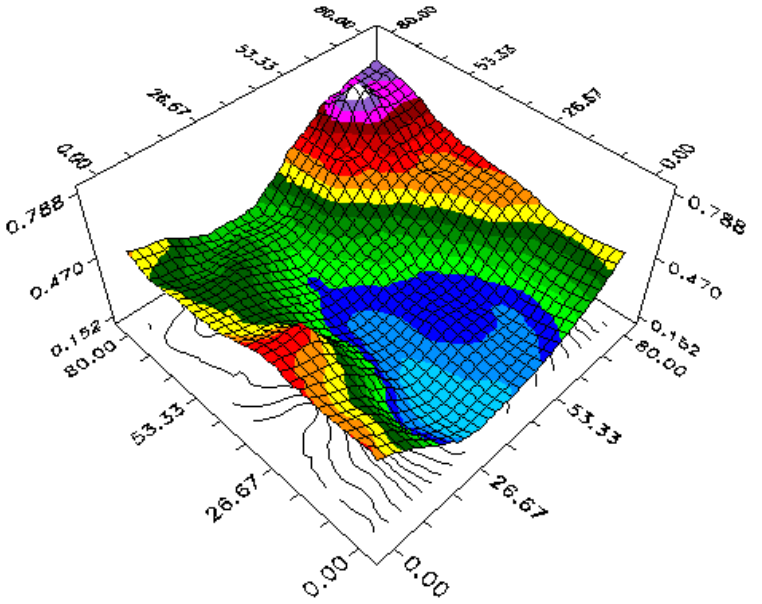
GS+ User's Guide
Version 10



GS+

GeoStatistics for the Environmental Sciences

G A M M A D E S I G N S O F T W A R E

Accessible Geostatistics for Everyday Science

GS+ ™

GeoStatistics for the Environmental Sciences

Version 10

Gamma Design Software, LLC
Plainwell, Michigan 49080
gammadesign.com

Copyright
© Copyright 1989-2015 Gamma Design Software, LLC. All Rights Reserved

ISBN number 0-9707410-0-6

Information in this document is subject to change without notice and does not represent a commitment on the part of Gamma Design Software. The software described is provided under a license agreement and may be used or copied only as specified in the agreement. No part of this document may be reproduced in any manner whatsoever without the express written permission of Gamma Design Software.

Gamma Design Software
P.O. Box 201
Plainwell, Michigan 49080
U.S.A.

Citation
The appropriate citation for this document is
Robertson, G.P. 2008. *GS$^+$: Geostatistics for the Environmental Sciences.* Gamma Design Software, Plainwell, Michigan USA.

Trademarks
Microsoft and Windows are trademarks or registered trademarks of Microsoft Corporation. Surfer is a registered trademark of Golden Software, Inc. ArcView and Arc/Info are registered trademarks of ESRI, Inc. Other brands and their products are trademarks or registered trademarks of their respective holders and should be noted as such. GS$^+$ is a trademark of Gamma Design Software.

January 2015

Table of Contents

Chapter 1
Introduction
- Overview....1
- System Requirements....2
- Installation and Activation....2
- Updates....11
- Licensing....12
- Single-User License Agreement....12

Chapter 2
Getting Started
- From Data to Maps: How to Proceed....14
- General Screen Layout....14
- Menus....15
- User Preferences....22
- Graph Settings....28
- Printing, Copying, and Exporting Graphs....34
- Printing, Copying, and Exporting Worksheet Data....34
- Using Older-Version GS+ Files....34

Chapter 3
Working with Data
- The Data Worksheet Window....36
- Column Assignments....38
- Covariate Values Warning....40
- Missing Values....40
- Data Filter Dialog....41
- Duplicate Values....42

Chapter 4
Importing Data from External Files
- File Import Dialog....43
- Importing Text Files....44
- Text Input File Formats
  - GS+....46
  - GeoEas....47
  - ArcView® XYZ....48
  - Surfer® XYZ....49
- Spreadsheet and Database Input Files....50
- Input File Formats (File Import Properties)....51

Viewing Files (File View Window) 52
Appending Data to an Existing Worksheet (Data Append Dialog) 52

Chapter 5
Summary Statistics
Z Variate Summary 53
Frequency Distributions
Frequency Distribution Histograms & Values 55
Cumulative Frequency Distribution & Values 57
Normal Probability Distribution & Values 58
X,Y Coordinates Summary 61
Coordinate Postings (Quantile Plots)
2-Dimensional Data 62
1-Dimensional Data 64
Defining Posting Intervals 65
Regression Analysis Tab 66
Cross-Variate Regression Window 67
Regression Values 68

Chapter 6
Semivariance Analysis
Overview 68
The Autocorrelation Window 70
Define Irregular Lag Class Intervals 74
Variograms Window 76
Semivariance Values Window 77
Isotropic Variogram Models 78
The Spherical Isotropic Model 81
The Exponential Isotropic Model 82
The Linear Isotropic Model 83
The Gaussian Isotropic Model 84
Anisotropic Variogram Models 85
The Spherical Anisotropic Model 87
The Exponential Anisotropic Model 88
The Linear Anisotropic Model 89
The Gaussian Anisotropic Model 89
Anisotropic Semivariance Surface (2-d Variogram Map) 91
Cross-Autocorrelation or Cross-Semivariance Analysis 94

Chapter 7
Other Autocorrelation Measures
Standardized Variograms 94
Madograms 96

Rodograms....97
Drift....98
Correlograms....99
Covariance Analysis....100
General Relative Variograms....101
Pairwise Relative Variograms....103
Moran's I Analysis....104
Fractal Analysis....106

Chapter 8
Variance Clouds and h-Scattergrams
Overview....107
Variance Clouds....107
Variance Cloud Pairs....111
h-Scattergrams....112
h-Scattergram Pairs....115

Chapter 9
Interpolation Basics
Overview....116
The Interpolation Window....117
Z Estimate Boundaries....121
Defining a Regular Interpolation Grid....121
Defining an Irregular Interpolation Grid....123
Polygon Outlines (Interpolation Masks)....125
Polygon Outline Map....127
Cross-Validation and Jackknife Analysis....128
Cross-Validation and Jackknife Values....130

Chapter 10
Conditional Simulation
Overview....131
Secondary Data for Simulations....133
Number of Simulations....134

Chapter 11
Kriging and Cokriging
Kriging....135
Cokriging....137

Chapter 12
Inverse Distance Weighting
Inverse (IDW) and Normal (NDW) Distance Weighting....140

Chapter 13
Interpolation Output File Formats
GS+ (.krg) Format 142
ArcInfo® or ArcView® (.asc) Format 143
Surfer® (.grd) Format 144
GSLib (.out) Format 145

Chapter 14
Mapping
The Mapping Window 146
Map Contour Intervals 150
3-d Maps 152
Standard Deviations 153
Rotation 154
2-d Maps 155
Standard Deviations 155
Sample Posting 156
Transects (1-d Maps) 157
Standard Deviation 157

Chapter 15
Bibliography 158

Chapter 16
Glossary 161

Index 169

Chapter 1
Introduction

GS+ is a **g**eo**s**tatistical Analysis program that allows you to readily measure and illustrate spatial relationships in geo-referenced data.

GS+ analyzes spatial data for autocorrelation and then uses this information to make optimal, statistically rigorous maps of the area sampled. The maps can be created in GS+ or in other mapping programs or geographic information systems.

When do I need geostatistics?

Geostatistics is useful when you need to make accurate, statistically rigorous maps created from incomplete data – which means whenever you make a map for a property that cannot be exhaustively sampled. Whether you are mapping oil deposits or plankton distributions, geostatistics allows you greater confidence in the interpolated values for the locations not actually sampled.

Statistics Provided by GS+

GS+ provides spatial autocorrelation analysis:

Semivariance analysis produces variograms and different types of variogram models, including isotropic and anisotropic variograms;

Anisotropic variogram maps make it easy to recognize anisotropy;

h-Scattergrams and Variance clouds provide an easy way to recognize data outliers; and

Many other types of autocorrelation measures, including Moran's I, fractals, correlograms, covariograms, madograms, rodograms, drift, standardized variograms, and general and pairwise relative variograms.

GS+ provides fast interpolation:

Various types of kriging provide optimal interpolation of a discrete point or an area around a sample point location;

Conditional simulation provides probability based interpolations and estimation error;

Cokriging provides optimal interpolation when you have only a few samples for the primary variate but many samples for an easier-to-measure, related covariate; and

Inverse distance weighting provides simple nearest neighbor interpolation based only on distance to nearby samples.

GS+ provides basic parametric statistics:

Sample means and variance;

Frequency distributions, probability distributions, and skewness and kurtosis measures for determining departures from normality; and

Quantile plots or coordinate maps show the distribution of sample values across the spatial domain;

Transformations for returning the data to normality; and

Regression analysis for covariates vs. primary variates.

System Requirements

GS+ requires a Windows-based PC running Windows XP or later.

Installation

To install GS+:

1) If GS+ is on a removable media such as a CD, insert the media into your drive and the install process should start automatically; if it does not, go to the drive for the media and double-click on SetupGSWin10.exe.

2) If you have downloaded GS+, then go to the folder into which you placed your download file and double-click on SetupGSWin10.exe.

3) The setup program will prompt you through the installation process. Follow the instructions on the screen. You will be prompted for a serial number and a license or activation code. These numbers can be found on the CD package or in your download confirmation email.

Note: You may be required to run the setup program as an administrator. If so, right-click on SetupGSWin10.exe and choose Run as administrator. For some configurations it may also be necessary to right-click the Properties menu and on the Compatibility tab, choose Run this program in Compatibility mode for Windows XP (Service Pack 3).

Activation

GS+ requires activation prior to first use once the evaluation period has expired. Activation is performed over the internet and requires a unique license code for each installation of GS+. Activation must be performed within several days of first installation (the activation screen provides the number of days remaining).

If you have already downloaded and installed an evaluation version of GS+, you do not need to reinstall GS+ once you have purchased a license. Instead choose "Activate GS+ now" from the start-up activation window (see below).

If you do not activate GS+ within the time period specified by the activation screen, GS+ will stop working. You may activate after this point without having to re-install the program.

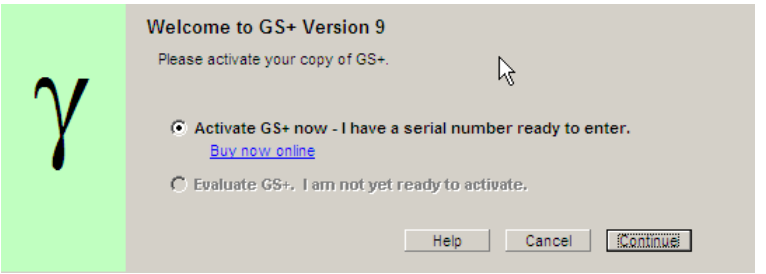
If the computer on which you want to activate GS+ is not connected to the internet, you may get a special unlocking code by logging into the activation server from a computer that is connected to the internet. See Advanced Options (below) for how to do this.

The license code required for activation is either on the CD sleeve or in the purchase confirmation email that you received.

Activation node-locks GS+ to a specific computer. Once your copy of GS+ has been activated, it will remain activated until you choose to deactivate it. You will want to deactivate it if you move GS+ to a different computer.

To Deactivate GS+ prior to moving the license to a different computer, use the Deactivate command on the Help menu. You do not need to deactivate GS+ prior to reinstalling Windows or reformatting your hard drive. GS+ may be deactivated and reactivated an unlimited number of times.

When you first launch GS+ the following activation screen will be displayed:


Welcome to GS+ Version 9
Please activate your copy of GS+.
Activate GS+ now - I have a serial number ready to enter.
Buy now online
Evaluate GS+. I am not yet ready to activate.
Help
Cancel
Continue

Acivation Options

1. Activate GS+ now – I have a serial number ready to enter.
   Choose this option if you have purchased GS+ and have a 15-digit license number (activation code) from the CD sleeve or from the purchase confirma-

tion email.

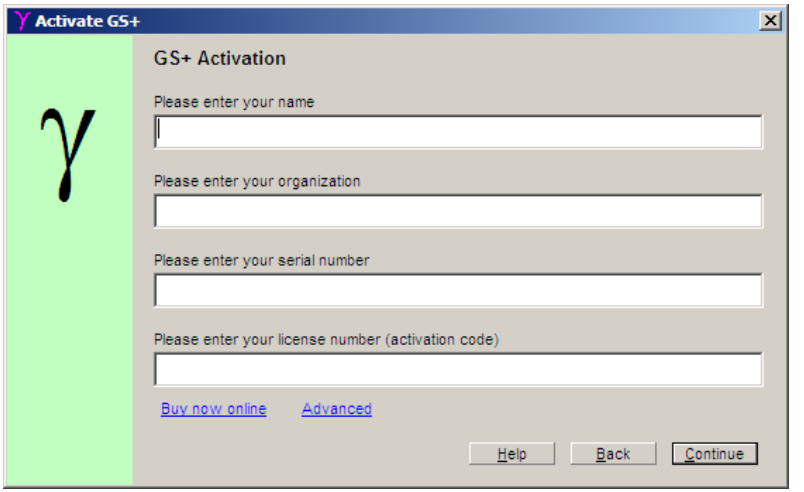
2. Evaluate GS+, I am not yet ready to activate.
   Choose this option if you want to remain in evaluation mode. You will have the number of days specified in the dialog narrative to purchase a license. If this option is not available (if it is disabled as in the illustration above), then you have run out of evaluation days and must purchase a license in order to continue.

Continue

Continue to the Activate GS+ window.

How to Activate

If you choose to Activate Now from the Activation Screen above, the following Activate GS+ window will be displayed:



Name, Organization

Enter your name and organization.

Serial Number

Enter the serial number for your copy of GS+. Find this number on the CD sleeve, or in the confirmation email you received after purchasing a downloaded copy of GS+. It will start with a letter.

License Number (activation code)

Enter the license number (also called activation code). This number is on the CD sleeve, or in the purchase confirmation email that you received.

Buy now online

Launch a web browser pointed to www.gammadesign.com that allows you to purchase a license.

Advanced

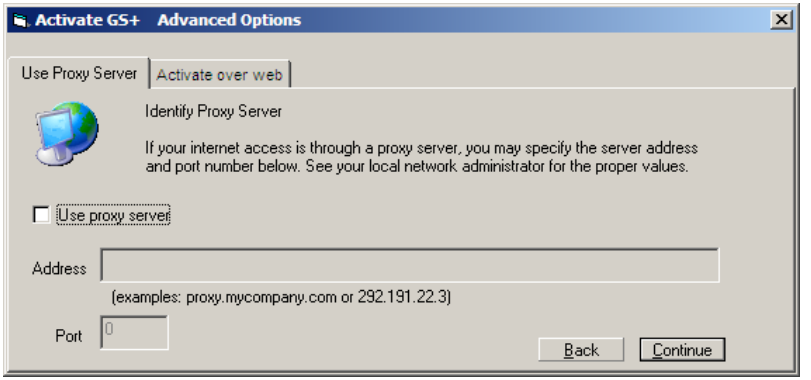
Bring up an advanced activation dialog that allows you to activate through a web browser on another computer, or to use a proxy server to connect to the internet.

Continue

Continue with activation.

Activation - Advanced Options

This dialog provides the option for providing a proxy server address for accessing the internet during activation or for activating over the web using this or another computer:


Activate GS+ Advanced Options
Use Proxy Server
Activate over web
Identify Proxy Server
If your internet access is through a proxy server, you may specify the server address and port number below. See your local network administrator for the proper values.
Use proxy server
Address
(examples: proxy.mycompany.com or 292.191.22.3)
Port
0
Back
Continue

Use Proxy Server

Check this box to connect to the internet using the proxy address and port specified. This is useful if your network settings do not allow direct connection, as is the case with some corporate and other servers.

Address

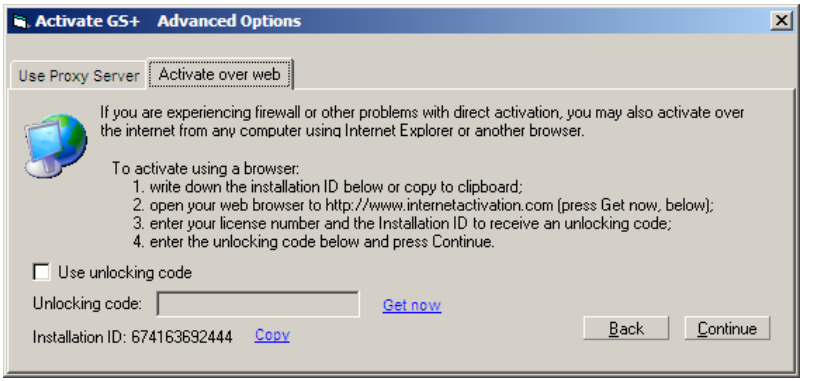
Specify the proxy address to use. Examples of valid address formats include proxy.mycompany.com and 192.191.22.33. Get the correct address from your network administrator.

Port

Specify a port to use for the proxy server. Get the correct port number from your network administrator.

Activate over web

Choose this tab to activate your copy of GS+ over the web rather than directly through the activation dialog. This is useful if you cannot otherwise connect to the internet activation site:


Activate GS+ Advanced Options
Use Proxy Server
Activate over web
If you are experiencing firewall or other problems with direct activation, you may also activate over the internet from any computer using Internet Explorer or another browser.
To activate using a browser:
1. write down the installation ID below or copy to clipboard;
2. open your web browser to http://www.internetactivation.com (press Get now, below);
3. enter your license number and the Installation ID to receive an unlocking code;
4. enter the unlocking code below and press Continue.
Use unlocking code
Unlocking code:
Get now
Installation ID: 674163692444
Copy
Back
Continue

Use unlocking code

Check this box to use an unlocking code to activate your copy of GS+. This is necessary if you cannot access the internet from this computer. In this case:

1. go to another computer and open a web browser to the address www.internetactivation.com
2. enter your license number (activation code) and the 12-digit Installation ID printed at the bottom of the dialog window (the Installation ID is different for every computer). Receive an unlocking code.
3. enter the unlocking code in the window below.

Unlocking code

Enter the unlocking code provided by the web site.

Get now

Launch a web browser pointed to www.internetactivation.com.

Copy

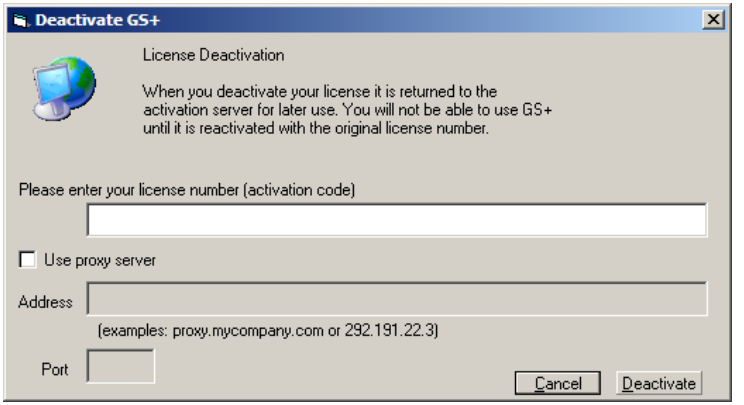
Copy the Installation ID to the Windows clipboard. This makes it easy to transfer into the browser window with an Edit | Paste command.

Continue

Return to the Activation Window. In the Activation Window you must still enter your License number (activation code) even with an Unlocking code present; the Unlocking code allows GS+ to activate without a direct connection to the internet.

Deactivation

You may transfer your activated GS+ license to another computer by first deactivating it on the current computer. After transferring your license, you will not be able to open GS+ again until it is reactivated (on the same or a different computer).You must be connected the internet in order to deactivate.


Deactivate GS+
License Deactivation
When you deactivate your license it is returned to the activation server for later use. You will not be able to use GS+ until it is reactivated with the original license number.
Please enter your license number (activation code)
Use proxy server
Address
(examples: proxy.mycompany.com or 292.191.22.3)
Port
Cancel
Deactivate

License Number (activation code)

Enter the license number (also called activation code). This number is on the CD sleeve, or in the purchase confirmation email that you received.

Use Proxy Server

Check this box to connect to the internet using the proxy address and port specified.

Address

Specify the proxy address to use. Valid formats include (a) proxy.mycompany.com or (b) 192.191.22.33. Get the correct address from your network administrator.

Port

Specify a port to use for the proxy server. Get the correct port number from your network administrator.

Deactivate

Deactivate the current license. Once you deactivate you may not run GS+ again on this computer without activating it again. There is no limit on the number of times you may deactivate and re-activate.

Updates

Maintenance updates are available free of charge to registered users. Update files are available by download only from http://www.gammadesign.com. The current version of GS+ can be checked from the Help menu as described in Chapter 2.

Licensing

Site Licensing

If you have reason to install GS+ on more than one computer in the same laboratory, or on a network that allows more than one user at a time to access the program, please contact Gamma Design for information about converting your single-user license to a laboratory group license (5-computer) or a single-building site license (unlimited computers in the same physical building). **It is a violation of your single-user license agreement if the program resides on more than one computer.**

Single-User (1-computer) License Agreement

Please read carefully; this is a legal End-User License Agreement (EULA) between you (the end user) and Gamma Design Software, LLC ("Gamma Design Software"). When you install this software on your computer you signal your agreement to be bound by the terms of this agreement, including the Software License and the Limited Warranty. If you do not agree to be bound by the terms of this agreement, do not install the software and return the package together with accompanying written material to Gamma Design Software at the address below for a full refund.

1. Gamma Design Software retains ownership of the GS+ program enclosed. **Gamma Design Software gives you (the end user) the right to use a single copy of GS+ on a single computer.** You may use GS+ on a network or file server ONLY if access is limited to one user at a time and you have the original copy of the documentation and program media.

   GS+ requires activation over the internet. You may transfer the license from one computer to another by deactivating the license on the computer on which it is currently installed and reactivating the license on another computer. **You do not have the right to activate or otherwise use GS+ on more than one computer, hard disk drive, or file server at a time.**

2. GS+ is owned by Gamma Design Software and is protected by United States copyright laws and international treaty provisions. GS+ must be treated like any other copyrighted material although you may either 1) transfer GS+ to a single hard disk drive so long as you keep the original copy for the purpose of backup, or 2) make one copy of GS+ for backup purposes. The written material accompanying GS+ may not

be copied.

3. GS+ may not be rented or leased, but may be permanently transferred if you keep no copies of any version of GS+ and the recipient agrees to the terms of this agreement.
4. You may not decompile, disassemble, or reverse-engineer GS+.
5. Gamma Design retains all rights not granted expressly herein. Nothing in this Agreement constitutes a waiver of Gamma Design's rights under any federal or state law.

Limited Warranty

1. Gamma Design Software warrants that GS+ will conform substantially to the accompanying written materials for a period of 1 year from the date of purchase, provided that GS+ is used on computer hardware and with the operating system for which it was designed.
2. Gamma Design Software disclaims all other warranties, either express or implied, including implied warranties of merchantability and fitness for a particular purpose. This applies to both the software itself and accompanying written materials. This limited warranty gives you specific legal rights; you may have others that vary from state to state.
3. Under no circumstances shall Gamma Design Software be liable for any damages whatsoever arising out of the use of or inability to use GS+, even if Gamma Design Software has been advised of the possibility of such damages. Such damages include but are not limited to damages for loss of profits or revenue, loss of use of the software, loss of data, the cost of recovering such software or data, the cost of substitute software, or claims by third parties. In no case shall Gamma Design Software be liable for more than the amount of the license fee, as set forth below. Some states do not allow the exclusion or limitation of liability for consequential or incidental damages, so this limitation may not apply to you.

User Remedies

1. Gamma Design Software's entire liability and your exclusive remedy shall be, at Gamma Design Software's discretion, either (1) refund of the purchase price or (b) replacement of the software that does not meet Gamma Design Software's limited warranty. In either case software must be returned to Gamma Design Software with a copy of the sales receipt. This warranty is void if failure has resulted from accident, abuse, or misapplication. Any replacement will be warranted for one year.
2. The warranties and remedies set forth above are exclusive and in lieu of all others, oral or written, express or implied. No Gamma Design Software distributor or employee is authorized to make any modification or addition to this warranty.

U.S. Government Restricted Rights

GS$^+$ software and documentation are provided with RESTRICTED AND LIMITED RIGHTS. Use, duplication, or disclosure by the Government is subject to restrictions as noted in subparagraph (c)(1)(ii) of The Rights in Technical Data and Computer Software clause at 52.227-7013. The manufacturer is Gamma Design Software, P.O. Box 201, Plainwell, MI 49080.

General

You must register your copy of GS$^+$ to be eligible for customer support and service. You may

do so on line at www.gammadesign.com or during the installation process. If you have questions about this agreement, write to Gamma Design Software, P.O. Box 201, Plainwell, MI 49080, U.S.A.

Chapter 2
Getting Started

From Data to Maps: How to Proceed

To make a map using GS$^+$:

- First, collect samples from known locations. The sample locations do not need to be evenly spaced or even to lie on a grid, you simply need to know their location in a Cartesian (x,y) coordinate system. Note that latitude and longitude are not Cartesian coordinates: if your data are in latitude / longitude coordinates you should first convert them to Cartesian units such as UTM; there are many calculators for doing this on the web;
- Second, bring the data into the GS$^+$ Data Worksheet; you can enter the data directly into the worksheet or import the data from a text file, spreadsheet, or another source; often the easiest way to import data is to cut-and-paste from the source spreadsheet or text file;
- Third, perform Semivariance Analysis to produce a variogram model of the autocorrelation present in the data;
- Fourth, use Kriging, Cokriging, or Conditional Simulation to produce an interpolation file that will contain optimal estimates of values at evenly-spaced intervals over the sample area; alternatively you can use Inverse Distance Weighting (IDW) but your interpolations will not be optimal; and
- Finally, draw a 3-d or 2-d Map of the property. This map will be an optimal, unbiased representation of the property over the area of interest. You can also produce a confidence map for the estimates, which will allow you to assess the statistical error associated with each estimated contour interval.

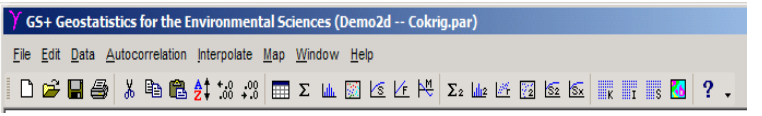
General Screen Layout

The main GS$^+$ window has a command menu at the top and holds each of the individual analysis windows that are currently open. The main windows are:

- The Data Worksheet Window
- The Data Summary Window
- The Autocorrelation Window
- The Interpolation Window
- The Map Window

Main Menu

The Main Menu presents access to the windows that provide GS+ analyses. Below the command menus are icons that represent short-cuts to many of these functions.

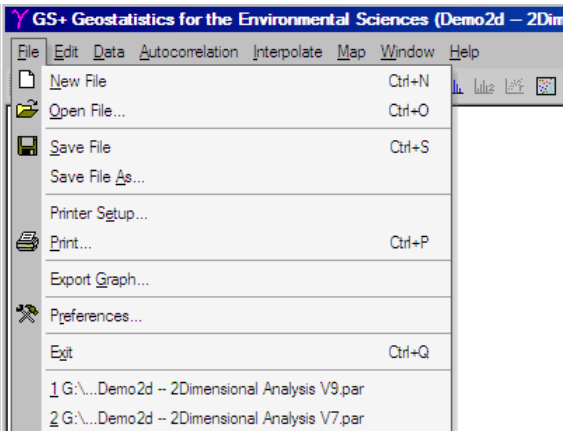

GS+ Geostatistics for the Environmental Sciences (Demo2d -- Cokrig.par)
File Edit Data Autocorrelation Interpolate Map Window Help

Commands and toolbar buttons that are shaded out are not selectable. For example, the print icon is not available when viewing a window without a graphic or worksheet to print. The six covariate icons for covariate functions are not available unless a covariate has been defined in the Data Worksheet window.

The tool bar is customizable via the Edit menu or by right-clicking on the main menu. You may add or remove shortcut icons for most analyses. It is also movable and reconfigurable; it can be dragged to any spot on the screen with the mouse.

The File Menu

The File Menu provides commands for saving and retrieving GS+ parameter files, used to store and retrieve analysis settings, and also commands for printing and setting user preferences.

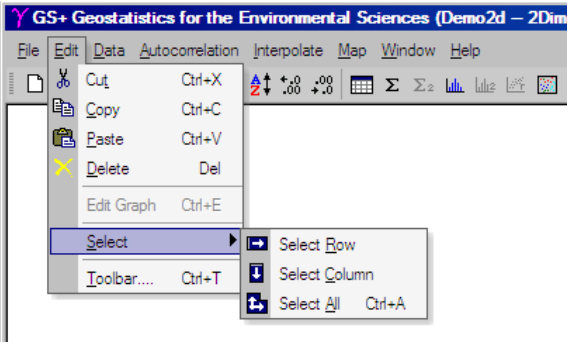


- New File – Clears existing analysis parameters.
- Open File – Allows the user to load an existing parameter file. To open a text data file, use the Import Command in the Worksheet window.
- Save File – Save the existing parameter file.
- Save File As – Save analysis parameters in a file to be named.

- Printer Setup – Make changes to the print format
- Print – Prints a spreadsheet or graph. You can also access this command by right-clicking on a graph or spreadsheet.
- Export Graph – Save graph to file; you can also access this command by right-clicking on a graph.
- Preferences – Set user preferences to be saved from session to session
- Exit – Exit GS+

The Edit Menu

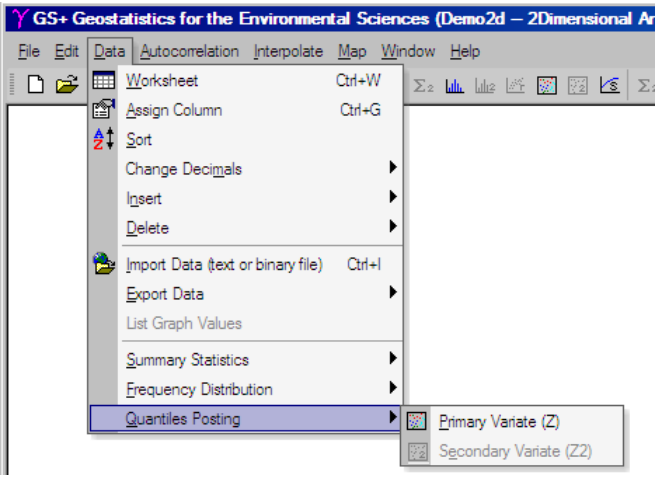
The Edit menu provides access to the cut-copy-paste-delete editing commands. These commands are available whenever the cursor is in an editable field within a particular window.



- Cut – remove selected material to the clipboard
- Copy – copy selected material (or graph, if graph window) to the clipboard. You can also access this command by right-clicking on a graph or worksheet.
- Paste – paste the clipboard into the selected area
- Delete – delete selected material
- Edit Graph – edit the graph in the active window. If there is not a graph in the active window, this command will be dimmed. You can also access this command by right-clicking on a graph.
- Select – Select a worksheet row or column. If the cursor is not in a worksheet this command will be dimmed.
- Toolbar – customize the toolbar to contain a variety of different tool shortcuts

The Data Menu

The Data menu provides access to the data worksheet and to summary statistics windows, and provides commands for importing and exporting data files and for manipulating data within the worksheet.

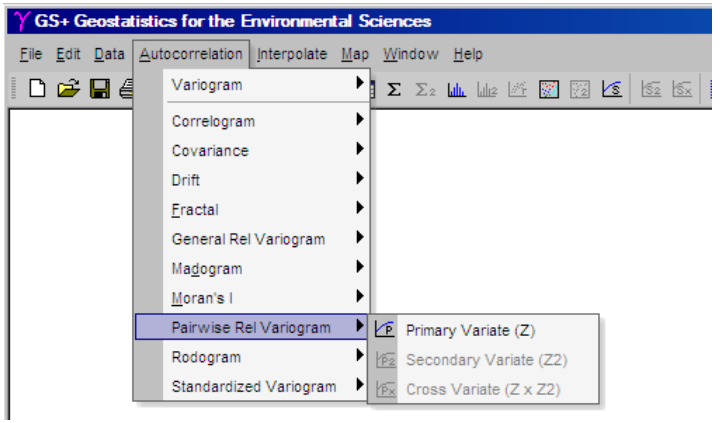


- WorkSheet – display the data worksheet window
- Assign Column – assign a variate to a data sheet column
- Sort – sort the highlighted column; the sort alternates between ascending and descending order. If the cursor is not in a worksheet cell the sort command will be dimmed.
- Change Decimals – increase or decrease the number of decimal places to show in a particular column; this has no effect on calculations, which are always performed in double-precision arithmetic. The cursor must be in an active worksheet window or these commands will be dimmed..
- Insert – insert a row or column into the active worksheet
- Delete – delete a row or column in the active worksheet
- Import Data – import a data file into the active worksheet
- Export Data – export the contents of either the active worksheet or the active data arrays to an external file. The active data arrays are the non-missing data in the ID, X-Coordinate, Y-Coordinate, Z-data, and Drift columns of the Data Worksheet window. If the Z data have been transformed (using a transformation command), the transformed z values will be printed in addition to the original z values in the Data Worksheet. Data can be exported as tab-separated text files (.txt), comma-separated-value text files (.csv), Excel spreadsheet files (.xls), html text files (.htm), or xml files (.xml).

- List Graph Values – display the graphed data. This command will be dimmed unless the active window contains a graph. You can also access this command by right-clicking on a graph. Data underlying maps can be displayed from the Map menu only from the map windows.
- Summary Statistics – display basic statistics for the primary variate (Z) and covariate (Z2) if defined.
- Frequency Distribution – display frequency distributions for the primary variate (Z) and covariate (Z2) if defined.
- Quartiles Posting – display a posting (map) for the primary variate (Z) and covariate (Z2) if defined.

The Autocorrelation Menu

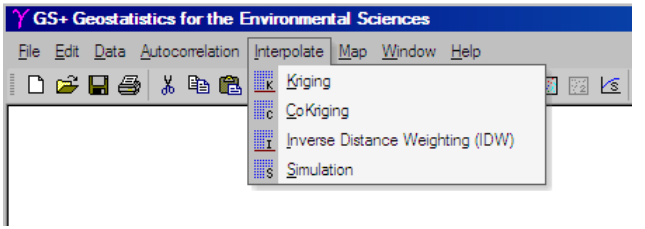
The Autocorrelation menu provides access to one of several types of autocorrelation analyses in GS$^+$:



- Variogram
- Correlogram
- Covariance Analysis
- Drift
- Fractal Analysis
- General Relative Variogram
- Madogram
- Moran's I Analysis
- Pairwise Relative Variogram
- Standardized Variogram

For each analysis you may choose to analyze the primary variate Z, the covariate Z2, or the cross-variate Z x Z2, although cross-variate analysis is not available for every autocorrelation measure. If Z or Z2 are undefined the menu command will be dimmed.

Interpolate Menu
The Interpolate menu provides access to one of the four major types of Interpolation provided by GS$^+$.



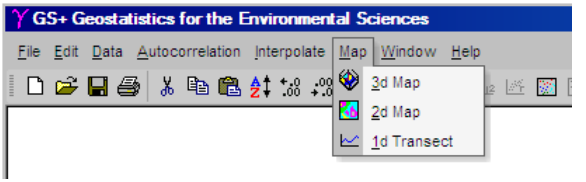
Kriging – display the Interpolation Window with access to Kriging

CoKriging – display the Interpolation Window with access to Cokriging

Inverse Distance Weighting (IDW) – display the Interpolation Window with access to Inverse Distance Weighting and Normal Distance Weighting interpolation

Simulation – display the Interpolation Window with access to Conditional Simulation

The Map Menu
The Map menu provides access to GS$^+$ mapping functions.


GS+ Geostatistics for the Environmental Sciences
File Edit Data Autocorrelation Interpolate Map Window Help
3d Map
2d Map
1d Transect

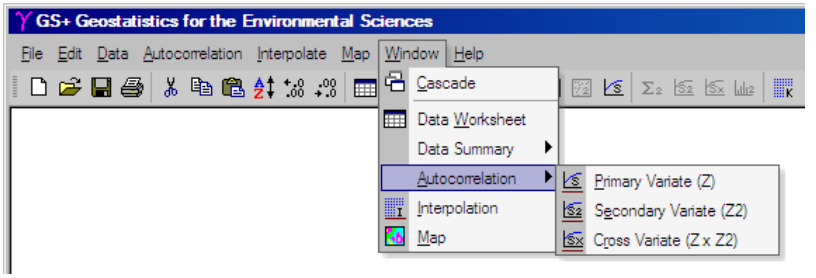
3-d Map – display the Mapping window with 3-d mapping selected

2-d Map – display the Mapping window with 2-d mapping selected

1-d Transect – display the Mapping window with 1-d mapping selected; if a 2-dimensional map file (one having both an x and a y coordinate) is currently selected, the 1-d option cannot be selected

Window Menu

The Window menu allows you to quickly move to open windows in GS+.



Cascade – rearrange all open windows

Data Worksheet – display data worksheet window

Data Summary – display data summary window for the primary or secondary variate

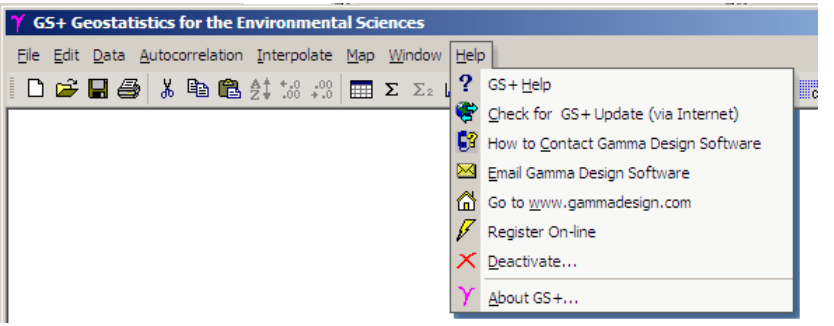
Autocorrelation Analysis – display autocorrelation windows for the Primary Variate (Z), Secondary Variate (Z2), or Cross Variate (Z x Z2).

Interpolation – display kriging, conditional simulation, and IDW analysis window

Map – display mapping window

The Help Menu

The Help menu provides access to GS+ help functions.


GS+ Geostatistics for the Environmental Sciences
File Edit Data Autocorrelation Interpolate Map Window Help
GS+ Help
Check for GS+ Update (via Internet)
How to Contact Gamma Design Software
Email Gamma Design Software
Go to www.gammadesign.com
Register On-line
Deactivate...
About GS+...

GS+ Help – display context-sensitive help topic.

Check for GS+ Update – send an automatic query to www.gammadesign.com to see if the version of GS+ currently running has been updated. If a newer version is available, you will be asked if you would like to connect to the proper web page for an update.

For this feature to work, your computer must have access to the internet. Communication is conducted using your browser. You can also check the GS+ update status manually by checking your program's version (available from the About GS+ screen) against the version displayed at www.gammadesign.com.

Email Gamma Design – send an email message to support@GammaDesign.com using your default email program. For this feature to work, you must have access to the internet and a default email program installed.

Go to www.gammadesign.com – connect to Gamma Design's home page through your normal internet provider using your default browser.

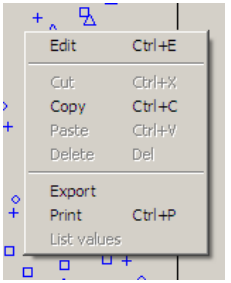
Register On-line – open a registration page to register with to Gamma Design Software.

Deactivate – transfer your GS+ license to the activation server. After transferring your license, you will not be able to open GS+ again until it is reactivated on this or another computer. You must be connected the internet in order to deactivate.

About GS+ – display title and copyright screen, and also display the current GS+ version number.

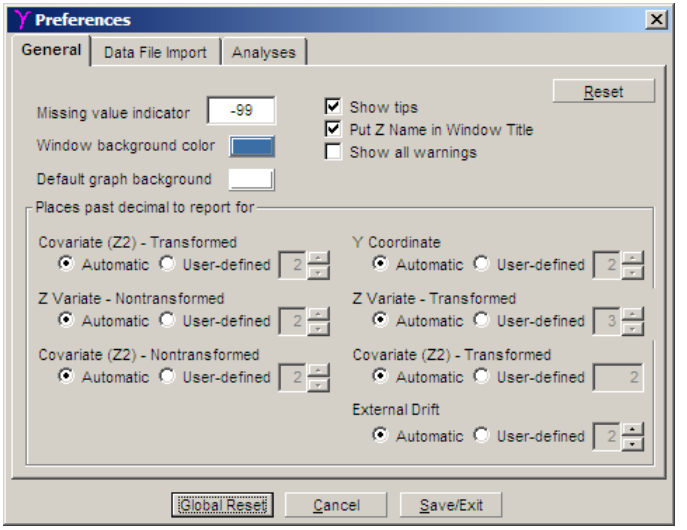
Right-Click (Context) Menus

When you right click over a graph or worksheet, a context menu will appear that will allow you to edit, print, export, or list the values in a graph, or print, or export worksheet values. Cut, Copy, Paste, and Delete functions are also available from the right-click menu.



User Preferences – General Tab

The Preferences dialog allows you to set user-default values for some GS+ behaviors. There are three categories available – General settings (as described here), settings for Data File Import, and settings for Analyses.



Reset

The **Reset** command returns all user-default values on this tab to original default values. To reset all values on all tabs, use the **Global Reset** command at the bot-

tom of the window.

Missing Value Indicator
Specify the value used by GS+ on output files and when importing data files to indicate that a value is missing. Missing values are ignored during data builds and subsequent analyses. This value can be overridden by values specified in individual dialog windows.

Window background color
Change the background color of the main GS+ window. The default color is the color of your desktop. Click on the color panel to change the color.

Default graph background color
Change the default Graph background color by clicking on the color panel. Default values can be overridden for individual graphs using the Graph Settings dialog.

Show tips
Display short explanations of commands and input boxes when mouse hovers over buttons and boxes.

Put Z Name in Window Title
Puts the name of the Z variate in the title bar of specific windows.

Show all warnings
Display warning messages even though you may have earlier elected to turn off warnings on some warning dialog screens where there is a "Do not show this message again" check box; when you choose to *Show all warnings* these messages are again shown.

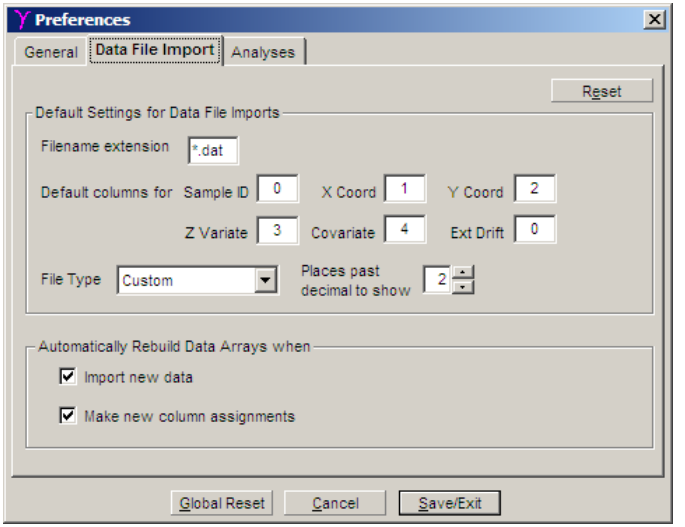
Places Past Decimal
For different types of variates, allow GS+ to format values automatically or specify the exact number of places past the decimal to report in windows and printouts. All calculations are performed on double-precision values regardless of the values requested here. These default values can be overridden by values on specific dialog windows such as the Field Assignment Dialog of the Data Worksheet Window.

Global Reset
Sets all user-default values on all tabs to original (GS+ - defined) default values.

Save/Exit
Close window and keep preference changes.

User Preferences - Data File Import Tab

The Preferences – Data File Import dialog window allows you to set user-default values for importing data files into the GS+ Data Worksheet.



Reset
The **Reset** command returns all user-default values on this tab to original default values. To reset all values on all tabs, use the **Global Reset** command at the bottom of the window.

Filename Extension
Default extension for the data file name specified when importing data files to the Data Worksheet.

Default Columns
When importing data files, these values indicate which columns or fields to assign initially to different variates. Column assignments can be changed in the Data Worksheet window.

Default Import File Type
When importing data files, this file type will be the default type specified. See Import File Types later for a description of specific types.

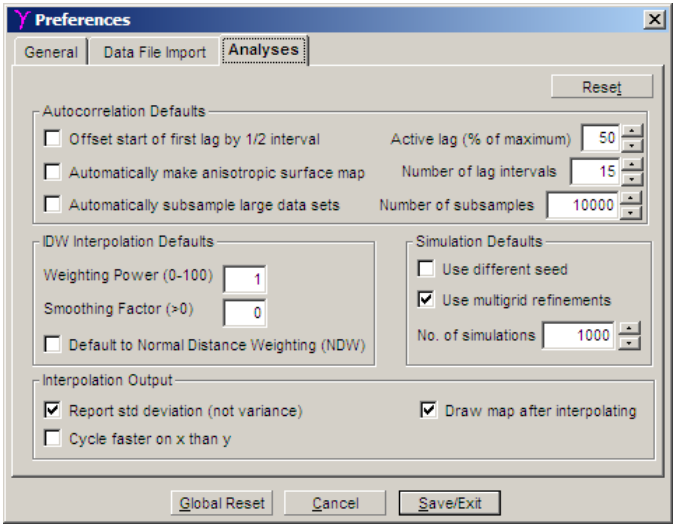
Places Past Decimal to Show
When importing data files, report this many places past the decimal in the Data Worksheet cells. This value is overridden when specifying places past the decimal in the Field (Column) Assignment Dialog of the Data Worksheet window.

Automatically Rebuild Data Arrays
When new data is imported and when columns are reassigned (e.g. when the X-Coordinate column is reassigned to column 3 from column 2), the internal data array used by GS$^+$ must be rebuilt. This can take some time if the array is big because duplicates must be checked, missing values tested, and autocorrelation boundaries calculated. The alternative to automatic recalculation is to press the Rebuild button on the Data Worksheet Window. Check or uncheck these boxes to turn off automatic array rebuilds.

Global Reset
Sets all user-default values on all tabs to original default values.

User Preferences – Analyses Tab

The Preferences – Analyses dialog window allows you to set user-default values for some analyses.



Reset
The **Reset** command returns all user-default values on this tab to original default values. To reset all values on all tabs, use the **Global Reset** command at the bot-

tom of the window.

Autocorrelation Defaults

- **Offset start of first lag by ½ interval** – with this check box GS$^+$ will (when automatically defining regular interval lag classes) shift the first lag class to ½ of its normal separation distance. This can provide better resolved variograms when there are sufficient pairs of points in shorter separation distance classes. The disadvantage is that if there are few pairs of points for the shortest distance classes when the first lag is not offset (a common problem), there will be even fewer pairs available with the first lag offset. If this is the case, there will be little if any improvement to the variogram.
- **Active lag (% of maximum)** – The Active Lag Distance specifies the range over which semivariance will be calculated. The default active lag is some percentage of the maximum lag, specified here. This is not likely to be the most appropriate active lag for your data but rather will provide a starting point for further exploration.
- **Automatically make anisotropic surface map** – the Surface Map command on the Semivariance Window will be automatically called after variograms are calculated. This will make the exploration of anisotropic relationships more automated but there is a performance penalty as creating a surface map requires further computation.
- **Number of lag intervals** – The active lag distance is divided into a number of different lag intervals classes for analysis. The number of classes chosen initially (by default) is specified here. During semivariance analysis the number of lag intervals will change as a function of the Active Lag divided by the Lag Class Distance Interval (see Autocorrelation Analysis, later).
- **Automatically subsample large data sets** – very large data sets can take hours of cpu time under certain conditions. Check this box to randomly subsample the data set during autocorrelation analysis only. When this box is checked and the data set being analyzed has more than the specified number of subsamples, only that number of subsamples are used in the autocorrelation analysis.

  Locations used in the analysis are randomly selected, and each subsampled location is compared against every other location in the data set regardless of whether the other location was also selected as a subsample – in this way there is no bias against small distance intervals due to subsampling. As an example of the time savings possible, 10,000 subsamples of a 160,000 point data set took 10 cpu minutes to produce essentially the same semivariogram model that otherwise took >8 cpu hours. Each of the 10,000 subsampled points was compared to the 159,999 other points to produce the variogram,

as opposed to when the full data set was sampled by comparing each of 160,000 points to 159,999 other points.

- **Number of subsamples** – specify the number of data records to use when subsampling large data sets.

IDW Interpolation Defaults

- **Weighting Power (0-100)** – this value provides the initial default weighting power for inverse distance weighting (IDW) interpolation. This value may be overridden from the Interpolate Window; see Chapter 12, IDW and NDW, for further information.
- **Smoothing Factor (>0)** – this value provides the initial default smoothing factor for inverse distance weighting (IDW) interpolation. This value may be overridden from the Interpolate Window; see Chapter 12, IDW and NDW, for further information.
- **Default to Normal Distance Weighting (NDW)** – sets the default IDW interpolation to Normal Distance Weighting (NDW). This value may be overridden from the Interpolate Window; see Chapter 12, IDW and NDW, for further information.

Simulation Defaults

- **Use different seed** – if this box is set each simulation will use a different random number seed by default. See Chapter 10, Conditional Simulation for further details.
- **Use multigrid refinements** – Checking this box forces the analysis to follow a stepwise procedure (when simulating interpolation nodes) by default. See Chapter 10, Conditional Simulation for further details.
- **No. of simulations** – specify the default number of simulations to perform. See Chapter 10, Conditional Simulation for further details.

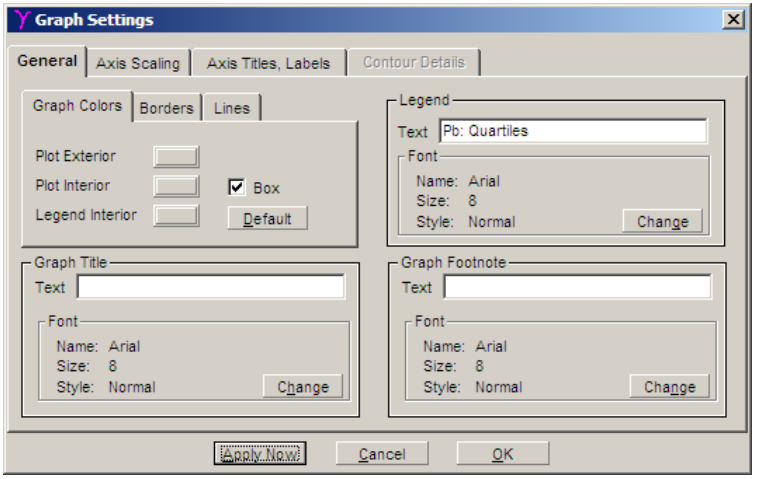
Interpolation Output

- **Draw map after interpolating** – check this box to automatically draw the map of interpolated values after interpolation.
- **Report std deviation (not variance)** – check this box to report standard deviation values rather than variance values for interpolation error terms. This is a default condition that can be changed in the Interpolation Window; see Chapter 9, Interpolation Basics, for further information.
- **Cycle faster on x than y** – ordinarily 2-dimensional interpolation results are written to output files with **y**-coordinate values cycling fastest (e.g. in the x,**y** order 10, **2**; 10, **3**; 10, **4**; 10, **5**; 20, **2**; 20, **3**; 20, **4**, 20, **5**); check this box to write results with **x**-coordinate values cycling faster (e.g. **10**, 2; **20**, 2; **10**, 3; **20**, 3; **10**, 4; **20**, 4; **10**, 5; **20**, 5).

Global Reset
Sets all user-default values on all tabs to original (GS+ defined) default values.

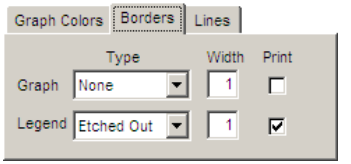
Graph Settings – General Tab

The Graph Settings Dialog appears when you edit graphs. The dialog will look slightly different depending on whether you are editing a bar graph, an x-y scatter graph (e.g. a variogram), or a 2-d or 3-d map. In the case of a 3-d map, for example, there will be a place for scaling and renaming the Z axis in addition to the X and Y axes. In addition to the General tab, there is also a tab for Axis Scaling, Axis Titles and Labels, and Contour Details (only available for maps).

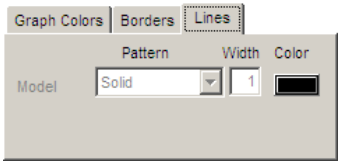


Graph Colors Tab
You may set background colors for three different parts of the graph. Click on the color bar to the right of the component name to bring up a Color Dialog Window that will allow you to change the color of that component; the color of the bar indicates the current color.

Graph Borders Tab
You may set border styles for the graph and for the legend (if shown). Width refers to the width of the border and Print allows you to choose to print or not print the border when printing:



Lines Tab
On graphs with lines (for example variogram models) you may specify line pattern, width, and color:



Graph Title
A title is text that appears at the top center of the graph area. To change the font of the title press **Change**, which will bring up a Font Dialog Window.

Graph Legend
The legend appears on maps and quantile postings to indicate the value ranges for different symbols or colors. Legend text appears at the top of the legend table. To change the font of the legend title and values press **Change**, which will bring up a Font Dialog Window.

Graph Footnote
A footnote is text that appears at the bottom left of the graph area. To change the font of the footnote press **Change**, which will bring up a Font Dialog Window. To have variogram model parameters printed as a footnote to variogram graphs, place in this field the exact phrase "(model results)" [do not include quotes]. This notation is placed in the field by default whenever a new variogram and model are calculated; you will need to remove it if you do not want variogram model parameters to appear as a graph footnote.

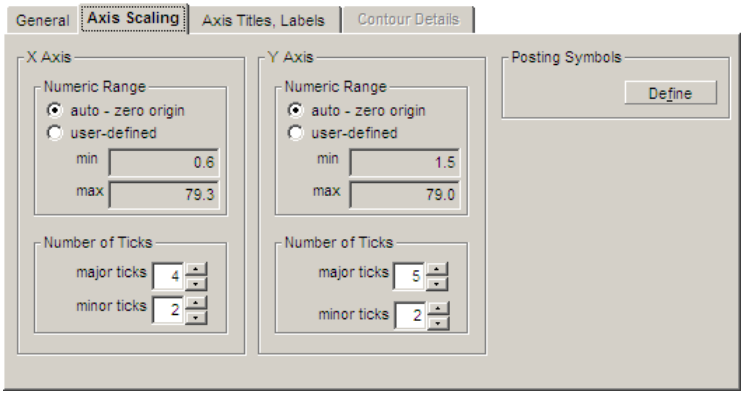
Apply Now
Press **Apply Now** to apply any changes made to the graph and keep the Graph Settings dialog window open.

Exit/Cancel
Press **Exit** to close the Graph Settings Dialog window. Any changes made since the last Apply Now command will be applied to the graph. Press **Cancel** to exit the Graph Settings Dialog without applying any changes since the last **Apply Now**

command.

Graph Settings - Axis Scaling Tab

The Axis Scaling tab of the Graph Settings Dialog Window allows you to specify how to scale the graph axes, e.g. how long or short an axis should be, and how the bars and symbols should look.



X Axis Range and Tick Marks

The X Axis range can be set to automatic or user-defined (manual). If the range is automatic and the lowest value in the graphed data set is greater than zero, then the axis range is set to a minimum value of zero and a maximum value of 10% greater than the highest value in the data set. If the range is automatic and the lowest value is less than zero, then the axis range minimum is set to 10% less than the lowest value.

The Number of Labels (bar graphs only; not shown above) refers to the number of values placed along the x-axis.

The Number of Ticks (x-y graphs and maps only) refers to the number of ticks along the x-axis. Major ticks are accompanied by labels (values); minor ticks are not labeled and appear between major ticks.

To set the appearance of the labels (font, precision, etc.) see the Axis Titles, Labels tab.

Y Axis Range and Tick Marks

The y-axis range is set identically to the x-axis range. Note that for maps of 1-dimensional transects there is no y-axis; the z-axis represents the vertical dimension.

Z Axis Range and Tick Marks

The z-axis range (on maps) is set identically to the X and Y Axes. Note that a z-axis is only present in maps and transects (1-d maps).

Bars

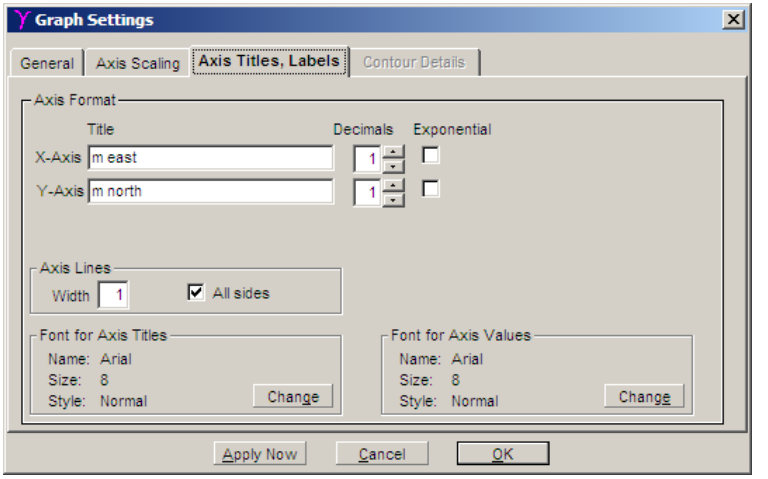
For bar graphs you may specify the number of bars to be plotted, their color, and pattern. The Bars section is not shown above.

Symbols

For x-y graphs (e.g. variograms) you may specify the type of symbol (open box, closed circle, etc.) as well as the size and color of the symbol. The Symbols section is not shown in the example screen above.

Graph Settings - Axis Titles, Labels Tab

The Axis Titles and Labels tab of the Graph Settings Dialog Window allows you to specify the text that accompanies each axis and to format the values that accompany the major tick marks.

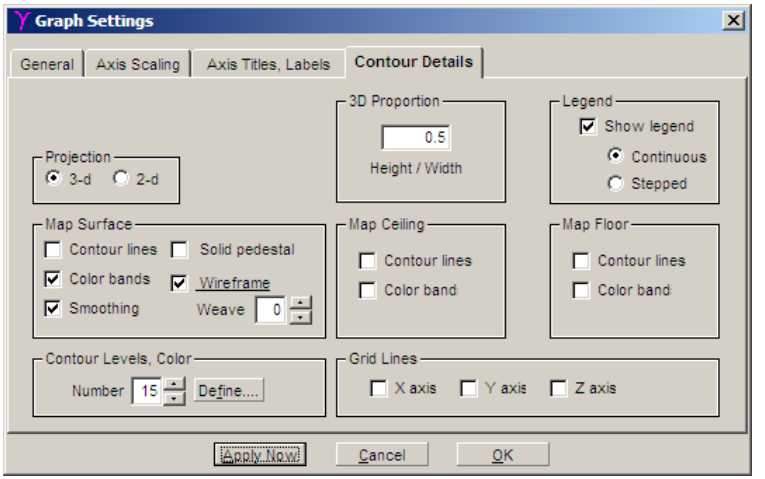


Axis Format

Use these boxes to set the axis titles and how axis values are formatted. Decimals refers to the number of places past the decimal to format axis values (e.g. 3.1415 has 4 places past the decimal); check exponential to format the axis value in scientific notation (e.g. 3.14E0). For 3-d maps, boxes for the Z-Axis are also provided as shown above; for x-y and bar graphs, there will be no boxes for the Z-Axis. The font for axis titles and values can be reset with one of the **Change** commands.

Graph Settings - Contour Details Tab

The Contour Details Settings tab of the Graph Settings Dialog Window allows you to edit map contour parameters. Some of these settings are identical to those that appear in the Map Window and changing values here will also change values there; for other maps (e.g. Anisotropic Surface Maps) this window is the only place where contour details can be set.



Projection

3-d – displays a 3-dimensional map of the data. The height of the map can be adjusted by changing the 3-d Proportion; the perspective can be adjusted with the Rotate (Mouse Action) command of the calling window

2-d – displays a flat, 2-dimensional map of the data

3-d Proportion

for 3-d projections, sets the height-to-width ratio of the map. For tall, skinny maps set this value to 2 or higher. For flatter maps, set to less than 0.5. This setting is not available for 2-d projections. This command will not affect the aspect ratio of the map, i.e. the x and y axes will remain proportional to one another

Map Surface

Contour lines – draws lines between contour intervals

Color bands – fills the space between contour lines with different colors; colors can be specified from the Define Contour dialog window by pressing Define Contour Levels, below

Smoothing – apply slight smoothing to the data to improve visualization

Solid pedestal – for a 3-d map fill in the area beneath the surface with a solid color.

Wireframe – for 3-d maps, drapes an x-y grid over the surface

Wireframe Weave – specifies the density of the wireframe grid. A weave of 0 puts a grid line at every data row and column; a weave of 1 skips one row/column, a weave of 2 skips 2, etc.

Contour Levels

Number – the number of contour levels to put on the map

Define – brings up the Map Contour Intervals dialog window that allows you to set break points for individual intervals and colors for contour bands (see Chapter 12 for a description of the Contour Intervals dialog)

Map Legend

Show legend – displays the contour legend next to the map

Continuous vs. Stepped – display legend as continuous scale with values next to break points (e.g. between color bands) or display legend as stepped format, with values next to boxed colors

Ceiling

Contour lines – project contour lines above the surface onto the ceiling of the plot (3-d maps only)

Color bands – project color bands onto the ceiling of the plot (3-d maps only)

Floor

Contour lines – project contour lines beneath the surface onto the floor of the plot (3-d maps only)

Color bands – project color bands onto the floor of the plot (3-d maps only)

Grid Lines

X axis – place a vertical grid line along the back walls of the 3-d plot x axis

Y axis – place a vertical grid line along the back wall of the 3-d plot y axis

Z axis – place horizontal grid lines along the back walls of 3-d plots

Printing, Copying, and Exporting Graphs

Graphs can be printed via the File | Print menu. If the active window contains a graph when the Print command is executed, the graph on that window will be sent to the printer. You may also right-click on the graph to access a Print command. A standard Windows Print dialog will appear.

Graphs can be copied to the Windows clipboard via the menu command Edit | Copy (or the control-C shortcut). The window containing the graph must be active (selected) prior to pressing the Copy command in order for the graph to be copied. You can also Copy graphs by right-clicking on the graph to bring up a context menu.

Graphs can be exported as graphic files via the menu command File | Export Graph. You have the option of saving graphs as enhanced metafiles (.emf), standard Windows metafiles (.wmf), bitmap files (.bmf), PNG files (.png), or jpeg files (.jpg). You can also export graphs by right-clicking on a graph to access an Export command.

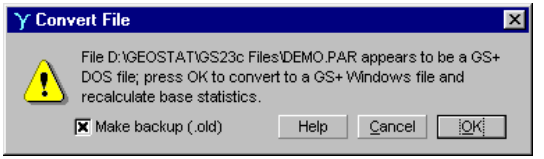
Printing, Copying, and Exporting Worksheet Data

Worksheet data can be printed, copied to the Windows clipboard, and exported to an external file in the same way that graphs can be. Right-click on a worksheet to bring up a context menu with these commands exposed. These tasks can also be performed via the main menu: printing can be performed through the File | Print menu, copying can be performed through the Edit | Copy menu, and exporting can be performed through the Data | Export menu (so long as the currently selected or active window contains a worksheet). If a block of cells is highlighted when printing, copying, and exporting is invoked, only that block will be acted upon.

Available export formats include text files with either tab-separated values (.txt) or comma-separated values (.csv), Excel spreadsheets (.xls), html files (.htm), and xml files (.xml). Column titles are exported along with data except when exporting a region of the worksheet (by first highlighting a block of cells). To export the entire worksheet without column titles use Edit | Select all before exporting.

Using Older-Version GS$^+$ Files: The Convert File Dialog

When loading a parameter (.par) file created with a legacy DOS version of GS$^+$, you will be queried to convert the DOS file to a GS$^+$ Windows file.


Convert File
File D:\GEOSTAT\GS23c Files\DEMO.PAR appears to be a GS+ DOS file; press OK to convert to a GS+ Windows file and recalculate base statistics.
Make backup (.old)
Help
Cancel
OK

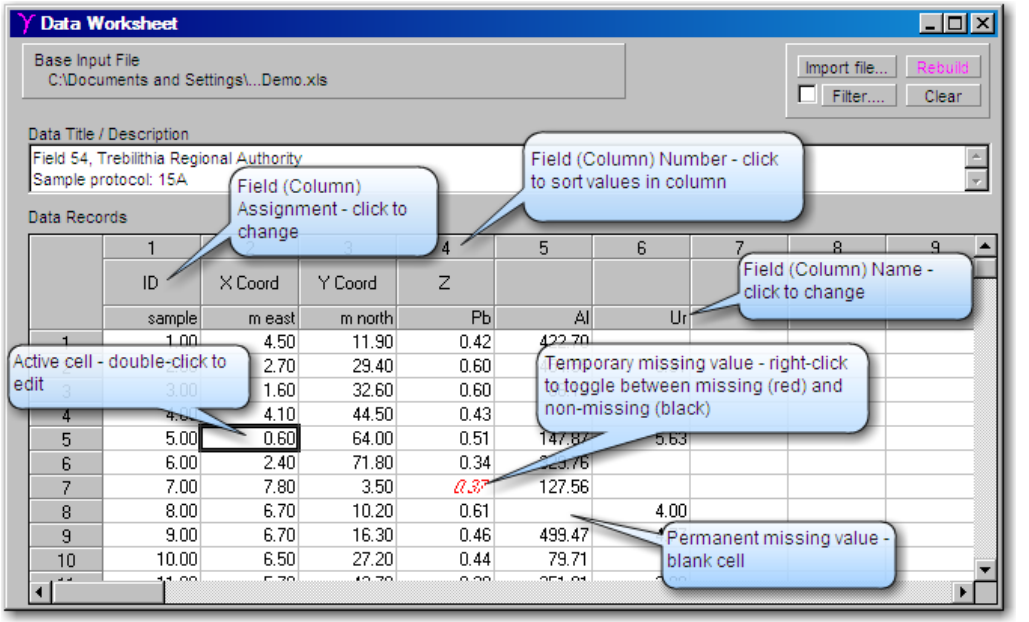
Make backup

Check this box to make a backup of the file before converting. The backup copy will have the same name as the original file but with the extension .old. E.G. a file named mydata.par will be copied to a file called mydata.old prior to the creation of the new, Windows-compatible mydata.par. If an error occurs during conversion, the original file will remain intact and the backup copy will not be made..

Chapter 3
Working with Data

The Data Worksheet Window

The Data Worksheet contains the data for GS+ analyses. Data can be entered manually or can be imported via the **Import file** command. Entered data can be edited, filtered (bounded), and can be temporarily or permanently deleted from subsequent analyses. Field assignments (assigning fields or columns to x-coordinate values, y-coordinate values, etc.) are made in the worksheet window by clicking on the second row.



Base Input File

The Base Input File is the external file from which worksheet data first were loaded – if none of the data were loaded from an external file then this field will be blank. To import data to the worksheet from an external file, press **Import** to bring up a File Import Dialog window. Note that the data in the worksheet may not be the same as data in the Base Input File if the worksheet data were changed after importing.

Import File

Press **Import File** to bring up a File Import dialog from which text and binary (e.g. worksheet) files can be imported into the worksheet (see Chapter 4).

Rebuild
The **Rebuild** command builds the data arrays on which all geostatistical analyses are based. This command is enabled whenever data records have been edited. The data arrays must be rebuilt prior to semivariance or other analysis whenever data have been changed. Sometimes this occurs automatically, such as when columns are reassigned. Other times, such as when individual cells are edited, you must rebuild the data arrays yourself by clicking on this button. When rebuilding is needed, the **Rebuild** command will become enabled and the color of the font will change to red. During rebuilding the data are checked for duplicate coordinate locations and for a sufficient number of valid records.

Note that rebuilding is only needed when column assignments change or when data within a column assigned to a coordinate or z-variate changes. In the screen above, changing a value within the Al column will not require arrays to be rebuilt because no analysis variates (X-coord, Y-coord, Z, or ID) are assigned to the Al column.

Filter
Press **Filter** to bring up a Filter Dialog that allows the data to be constrained to a particular range; data outside of the specified range are treated as temporary missing data, i.e. excluded from subsequent analyses. The filter check box turns filtering on and off.

Clear
Press **Clear** to empty the data worksheet and reset all analysis windows. Has the same effect as the **File – New** menu command.

Data Title / Description
Any text information desired can be entered in this field. When importing text files, the “header” records in the file – the records that appear prior to the data records – are automatically read into this box. If specified in the text file import window, these records can also contain variate names that appear as data column titles. The first line of the Data Description becomes the default graph and worksheet title. You can use this box for storing comments about the parameter file or analysis.

Data Records
The first row of the data worksheet holds the field or column numbers. You can click on a cell in this row to sort the contents of the column in ascending or descending order; each time you click the order is alternated. All row contents are sorted when you sort a particular field or column.

The second row of the data worksheet specifies the **Field or Variate Assignments**, i.e. which field or column contains the X-Coordinate Data, Y-Coordinate Data, Z-Variate Data, or Sample ID Data. These assignments can be changed by clicking within the top row, which will bring up a Field Assignment Dialog window.

The third row of the data worksheet specifies the user-supplied **Field or Variate Names** for the various data fields or columns. To enter or edit names, click on the

cell to be edited. When data is imported from a text or external worksheet file, variate names can be read from the header records.

The **data area** of the worksheet contain data for each variate. To enter or edit data in any given cell, double-click on that cell. To enter a formula rather than a value, begin the cell with an "=" sign.

The data worksheet has a capacity for 2 billion records and up to 64 columns.

To **sort** the datasheet, click on the topmost row of the column. Sorts will alternate between ascending and descending order. Or use the Sort command in the Data Menu.

To change the number of **decimal places** to show for any given column, use the Decrease Decimals or Increase Decimals command in the Data Menu, or click on the respective tool icon in the menu bar. If in the Preferences Window you have specified an automatic format for decimals to report, changing the number of decimal places for the X, Y, or Z columns will change the number of decimal places used in subsequent worksheets and graphs.

To change the width of any column , move the cursor to the top of the column and use the mouse to stretch or contract the column margins.

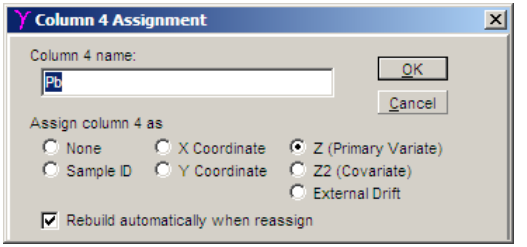
To temporarily delete a cell from analyses, change it to a **Temporary Missing Value** with a click of the right mouse button (its color turns red and the font becomes italicized; see the value *0.37* in the diagram above). Another right-click restores it to the worksheet (its color will return to black and font to normal).

To change a value to a **Permanent Missing Value**, delete its contents by highlighting it and pressing the Delete key (note the empty cells in the figure above).

To insert or delete a row or column in the worksheet use the Edit | Insert command or Delete on the menu bar.

Column Assignments

In the Data Worksheet window, you may specify which field or column to associate with the Sample ID, X-Coordinate, Y-Coordinate, the Z variate, the Z2 variate, or External Drift by clicking on the second cell from the top in any column. This will bring up the Column Assignment dialog:



Column name
The name of the column or variate name.

Assign column as

None – make column unassigned.

Sample ID – the specified column contains Sample Number or Sample ID information. This can be text or numeric data.

X Coordinate – the specified column contains values for the X-Coordinate location. If you choose a column that is already assigned, the other variate's column will switch with the original X Coordinate column.

Y Coordinate – the specified column contains values for the Y-Coordinate location. If you choose a column that is already assigned, the other variate's column will switch with the original Y Coordinate column.

Z (Primary Variate) – the specified column contains values for the Z variate. If you choose a column that is already assigned, the other variate's column will switch with the original Z column.

Z2 (Covariate) – the specified column contains values for the Covariate. The Covariate is used in cokriging. If you choose a column that is already assigned to another variate, the other variate's column will switch with the original Z2 column.

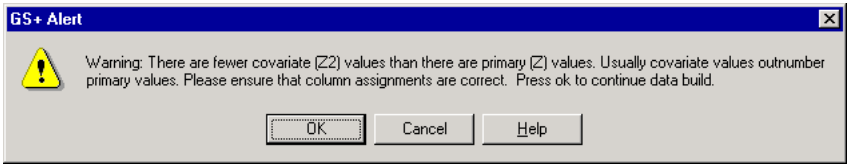
External Drift – the specified column contains the external drift data. If you choose a column that is already assigned to another variate, the other variate's column will switch with the original external drift column.

Rebuild automatically when reassign
Rebuild the internal worksheet when window closes. This is equivalent to pressing the Rebuild button on the Data Worksheet window.

Covariate Values Warning

In cokriging there are usually more Z2 covariate values than primary (Z) values. If your column (field) assignments for Z and Z2 result in data arrays with more Z values than Z2 values then you will get the warning notice below.

You may continue the analysis with more Z than Z2 values, but there is no advantage to doing so – in cokriging Z values without corresponding Z2 values are treated as missing. Likewise, there is no advantage to cokriging when you have just as many Z2 covariate values as you have Z values.


GS+ Alert
Warning: There are fewer covariate (Z2) values than there are primary (Z) values. Usually covariate values outnumber primary values. Please ensure that column assignments are correct. Press ok to continue data build.
OK
Cancel
Help

Missing Values

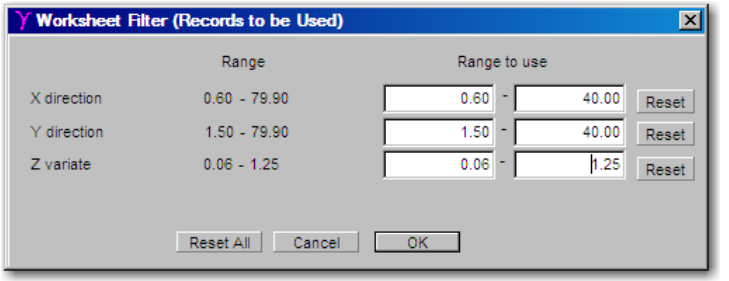
Data in the Worksheet that are marked as Missing are ignored during analyses. Permanent missing values appear as blank cells and temporary missing values appear in red. You can use the right mouse button to make cells temporarily missing and vice versa.

When importing or exporting files special placeholders (values or symbols) can be used to indicate missing values in the incoming or outgoing data file. These placeholders are specified by the user in the Preferences window, or during file imports by a value or symbol specified in the File Import Properties window. In GS$^+$, the default missing value indicator is the numeric value -99.0 but this can be changed in the User Preferences window (Chapter 2).

Missing values appear in output files when a value cannot be interpolated because the location appears in an exclusive polygon or because numerical limitations disallow its computation (such as when a variogram model is inappropriately used during kriging).

Data Filter Dialog

The data within a worksheet can be collectively filtered or bounded using the Filter command on the Data Worksheet window. With this command all records are scanned and if a record falls outside of the specified Range to Use, the variate outside of its range will be treated as missing values during the data build.



Range to Use

In these fields specify how to constrain the data in the Worksheet to a particular range. In the example shown, only those records for which the X direction field is within the range of 0.60 to 40.00 will be included in analyses even though there are records across the entire range of 0.60 – 79.90 in the worksheet. *Although Worksheet values are treated as missing during data builds, they are not marked as missing in the Worksheet.*

Reset and Reset All

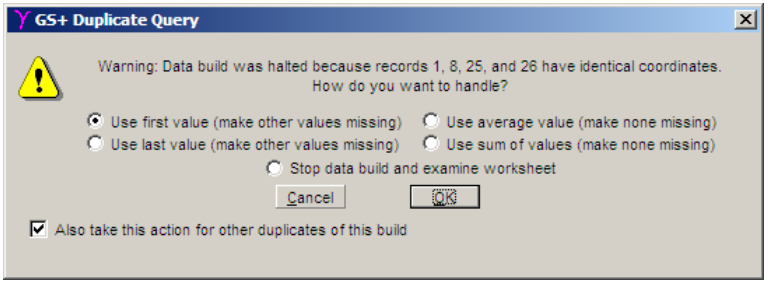
Reset Range to Use to the actual data range for a given coordinate or variate. Data ranges can be set individually or (with the **Reset All** command) all at once.

OK

Press OK to accept the ranges specified. Note that for the filter to take effect you must check the filter box on the Worksheet window. Press Cancel to exit the dialog without saving changes.

Duplicate Values

When worksheet data are rebuilt into data arrays, GS+ checks each record against every other record to check that no duplicate coordinates are present. If one or more records are found to be duplicates, a warning screen will appear and you will be asked how to resolve the duplicate:


GS+ Duplicate Query
Warning: Data build was halted because records 1, 8, 25, and 26 have identical coordinates.
How do you want to handle?
Use first value (make other values missing)
Use average value (make none missing)
Use last value (make other values missing)
Use sum of values (make none missing)
Stop data build and examine worksheet
Cancel
OK
Also take this action for other duplicates of this build

Use first value (make other values missing)
This option includes the first record encountered (record 1 in the case described) in the analysis and marks other duplicates in the worksheet (records 8, 25, and 26) as temporarily missing values. So long as these values remain marked as missing, on subsequent worksheet rebuilds this Duplicate Query warning will not occur.

Use last value (make other values missing)
This option is the same as the first option above except that the last record encountered is used in the analysis (record 26 in the case described) rather than the first.

Use average value (make none missing)
This option directs GS+ to use the average of the duplicate records. It does not mark any values in the worksheet missing, which means that the next time the worksheet is rebuilt this Duplicate Query warning will occur again.

Stop data build and examine worksheet
Specifying this option will cause the data build to halt and will display the first duplicate record of the data worksheet. Any other duplicates already found during the data build will be reset to non-missing will be returned to their pre-build state.

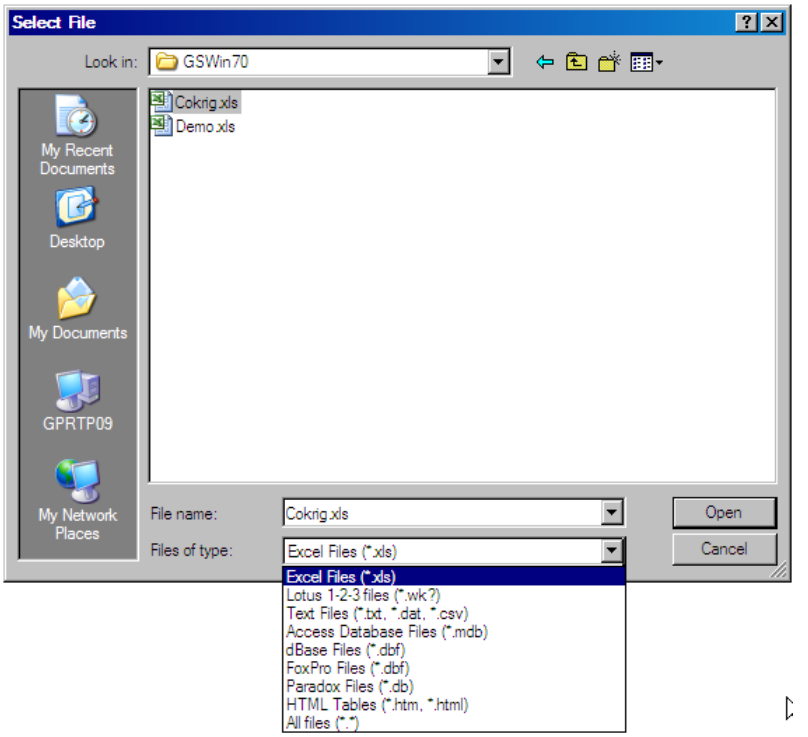
Also take this action for other duplicates of this build
Checking this box causes other duplicates encountered to be treated the same as this set of duplicates. For data sets with many duplicate values, this option saves time by keeping this dialog window from displaying again during the present build.

Cancel
Cancel the data build. Any duplicates set to missing values during this build will be reset to nonmissing.

Chapter 4
Importing Data from External Files

File Import Dialog

A number of different file types can be imported into GS+. These include spreadsheets (e.g. Microsoft Excel formats), databases (e.g. Microsoft Access formats), HTML (web) pages, and text files. Files are imported via the Import command within the Data Worksheet Window, which brings up the dialog box below. From this dialog you choose the file and the format (as illustrated in the pull-down list box at the bottom of the menu).



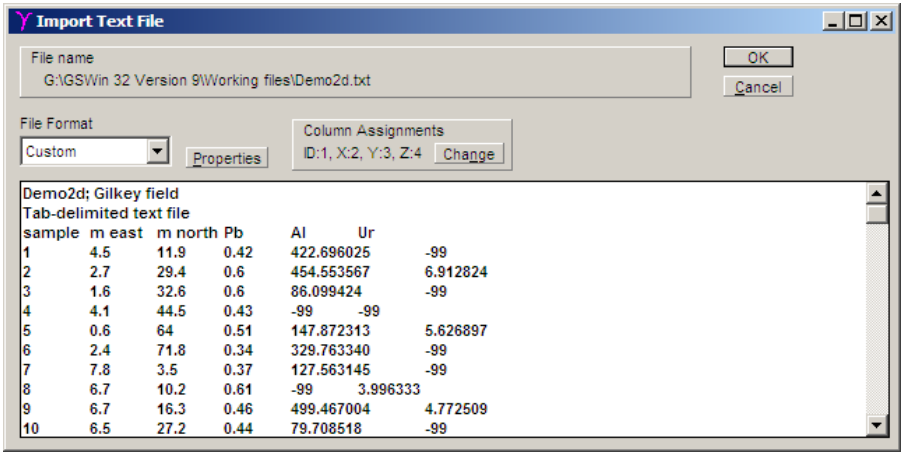
Open

Press Open to open the identified file. The window that next appears will depend on the type of file to be opened. A text file will bring up the Import Text File dialog whereas spreadsheet, database, and HTML formats will bring up the Import Spreadsheet dialog.

Importing Text Files

Choosing to import text files from the File Import Dialog brings up the window below. The contents of the file are displayed in the large preview area. When the OK button is pressed the contents of the file are read into the GS+ Worksheet according to rules specified by the chosen File Format. These rules define how many rows of header data precede the data rows (2, in the example below), how missing values are identified (with the number -99 in the example below), how column names are separated (with commas, below), and how data fields (columns) are separated from one another within data records (with spaces in this example).

A different dialog window is available for importing spreadsheet, database, and HTML files (described later).



File Format

A variety of text file types can be imported into GS+, including formats defined by the user. Each type has its own manner for separating fields within data records, for handling missing values, for allowing header records, and for specifying names of fields (column titles) within the file. Predefined input format types include the following; examples of files appear in the next section:

- GS+ format , in which fields are separated by spaces (free format) or tabs, missing values are indicated with the placeholder -99, the number of header records is automatically detected, and column titles (variate names) appear on the 2nd record separated by commas. An example appears in the next section.
- GeoEas format, in which fields are separated with commas or spaces, there are no missing value indicators, the number of header records is specified on the 2nd record of the file, and column titles appear as individual records following this second record. An example appears in the next section.

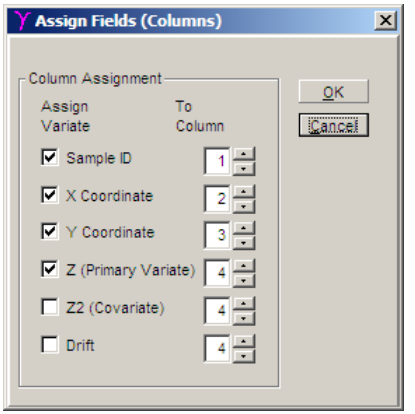
- Surfer® XYZ format, in which fields are separated by spaces (free format), missing values are indicated by blank fields, and the first record in the file is a header record in which column titles (variate names) appear as fields separated by spaces. Note that the Surfer XYZ format also allows fields to be separated by commas, which should be specified separately as described below. Note that this format is not the same as the Surfer Grid file format that can be used for Krig output files or Map input files. An example appears in the next section.
- ArcView® XYZ format, in which fields are separated with commas, missing values are indicated by placeholder commas, there is a single header record, and within the header record variate names are separated by quotes and commas. An example appears in the next section.
- Custom, in which any of these format specifications can be changed or customized as specified in File Import Properties as described in the next section.

Properties

The Properties Command brings up a File Import Properties dialog window (described below) within which you can specify the rules for GS+ file imports.

Change Column Assignments

The Change command brings up a Field or Column Assignment window from which you can assign variates (e.g. X coordinate) to columns or fields in the data file (see below). Columns can also be assigned later from the Data Worksheet window.



Text Input File Formats: GS+ Format

The standard GS+ input file format is comprised of header records and data records:

- Data records are space-delimited or tab-delimited XYZ type data. This means that each data record contains at least 3 fields: an x-coordinate location, a y-coordinate location, and the value for at least one z variate measured at that x-y location (single-dimension transects will have only x-coordinate and z variate data values). Additional fields can hold a sample ID value and multiple z-variates for a particular x-y location.
- Header records precede the data records and contain whatever text information about the file that the user feels is useful. There can be any number of header records; for this format GS+ determines the number of header records automatically, which means that data records start with the first all-numeric record. The last header record can contain column titles (variate names), separated by commas.
- Missing values are denoted by the number -99.
- Any of these parameters (field delimiters, number of header records, missing value indicators, etc.) can be changed to a custom format from the Import Data File window.
- The following listing is the first 11 records of a standard GS+ input file that has fields for sample ID, x-coordinate, y-coordinate, and two z variates. Note the variate names in the second record, and two missing values in the last column:

```
File Demo2-d.dat;
sample,  m east,  m north,  Pb,      Al
       1      4.5      11.9     0.42     0.42
       2      2.7      29.4      0.6     0.45
       3      1.6      32.6      0.6     0.08
       4      4.1      44.5     0.43   -99.00
       5      0.6        64     0.51     0.14
       6      2.4      71.8     0.34     0.32
       7      7.8       3.5     0.37     0.12
       8      6.7      10.2     0.61   -99.00
       9      6.7      16.3     0.46     0.49
```

Text Input File Formats: GeoEas Format

The standard GeoEas input file format is comprised of header records and data records :

- Data records are comma or space delimited XYZ type data. This means that each data record contains at least 3 fields: an x-coordinate location, a y-coordinate location, and the value for at least one z variate measured at that x-y location (single-dimension transects will have only x-coordinate and z variate values). Additional fields can hold a sample ID value and multiple z-variates for a particular x-y location.
- Header records precede the data records and contain specific information about the data records:

  Record 1 contains text of the user's choice, usually a data set title or file name.

  Record 2 contains the number of fields (values) in each data record.

  Record 3 contains the name of the first field.

  Records 4+ contain the names of the second, third, etc. fields (one name per record)
- There are no missing value indicators in the standard GeoEas format
- Any of these parameters (field delimiters, number of header records, missing value indicators, etc.) can be changed to a custom format from the Import Text File window.
- The following listing is the first 11 records of a standard GeoEas input file that has fields for sample ID, x-coordinate, y-coordinate, and one z variate. Note the four variate names in records 3-6:

```
File for field 55
4
ID
X meters
y meters
Z (mm)
13,-14036,-3097,11
14,-13621,-1266,25
22,-12384,-911,79
23,-121276,978,91
24,-12674,190,14
```

Text Input File Formats: ArcView® XYZ Format

The ArcView® input file format is comprised of header records and data records :

- Data records are comma-delimited XYZ type data. This means that each data record contains at least 3 fields: x-coordinate location, a y-coordinate location, and the value for at least one z variate measured at that x-y location (single-dimension transects will have only x-coordinate and z variate data values). Additional fields can hold other variates for that location, e.g. sample ID, other measured z-variates. The ArcView XYZ input format is not the same as the ArcInfo Ascii Raster File Format, which is a Kriging output format that can be read directly by Arcinfo® or Arcview® GIS software.
- A single header record precedes the data records and contains field (variate) names for the data record fields. Names are within quotes and comma-delimited.
- Missing values are indicated by blank fields.
- Any of these parameters (field delimiters, number of header records, missing value indicators, etc.) can be changed to a custom format from the File Import Properties window.
- The following listing is the first 9 records of a standard ArcView® input file that has fields for an x-coordinate, y-coordinate, and one z variate. Note the three variate names in record 1 and the missing values in records 5 and 9:

```
"X data","Y data","Z data"
4.5,11.9,0.42
2.7,29.4,0.45
1.6,32.6,0.08
4.1,,0.65
0.6, 64,0.14
2.4, 71.8,0.32
7.8, 3.5,0.12
6.7,10.2,
6.7,16.3, 0.49
```

Text Input File Formats: Surfer® XYZ Format

The standard Surfer® input file format is comprised of header records and data records :

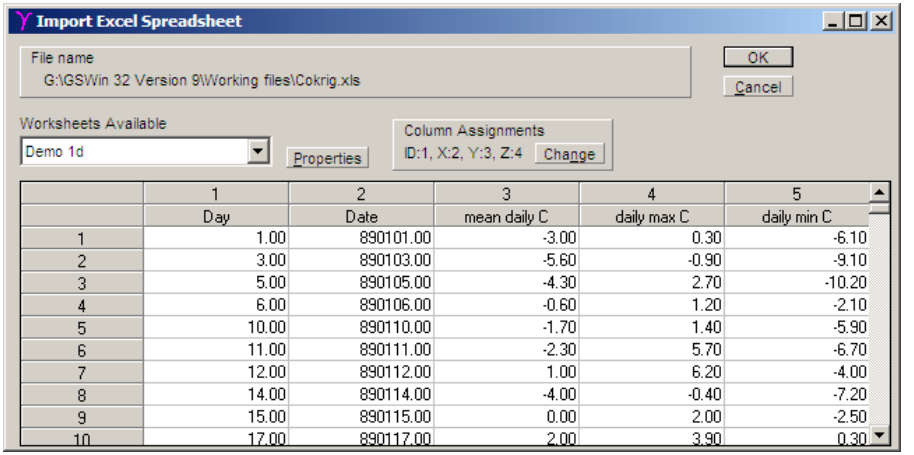
- Data records are space-delimited XYZ type data. This means that each data record contains at least 3 fields: an x-coordinate location, a y-coordinate location, and the value for at least one z variate measured at that x-y location (single-dimension transects will have only x-coordinate and z variate data values). Additional fields can hold other variates for that location, e.g. sample ID, other measured z-variates.
- A single header record precedes the data records and contains field (variate) names or column titles for the data record fields. Names are space-delimited so they must be single words (e.g. "mEast mNorth Nitrate") in order that they be properly assigned to their columns. You can allow names to be delimited by commas or other characters by changing this to a Custom Format in the File Import Properties window (e.g. "meters East, meters North, Nitrate (ug/L)").
- Missing values are indicated by blank fields. For files where there is more than one z-variate per record, a missing value for any field in the record means that the entire record will be treated as missing. (To avoid this problem use comma-delimited data records.)
- Any of these parameters (field delimiters, number of header records, missing value indicators, etc.) can be changed to a custom format from the Import Text File window. Note that some Surfer® files are comma-delimited rather than space-delimited.
- The following listing is the first 9 records of a standard Surfer® input file that has fields for an x-coordinate, a y-coordinate, and one z variate. Note the three variate names in record 1 and the missing values in records 5 and 9:

```
Xdata   Ydata   Zdata
4.5     11.9    0.42
2.7     29.4    0.45
1.6     32.6    0.08
4.1     44.5
0.6       64    0.14
2.4     71.8    0.32
7.8      3.5    0.12
6.7     10.2
6.7     16.3    0.49
```

Spreadsheet and Database Input Files

Choosing to import a spreadsheet, database, or HTML files from the File Import Dialog brings up a window similar to the one for Excel spreadsheets below. Multiple worksheets or tables within the larger spreadsheet workbook or database are available through a pull-down listbox. The contents of the spreadsheet or data table are displayed in the large preview area. When the OK button is pressed the contents of the file are read into the GS+ Worksheet according to rules specified by the **Properties** command. A different dialog is used to import text files as described earlier.

When importing spreadsheets, the first row of the imported spreadsheet should hold the variate names that appear in the GS+ column headings ("Day" in column 1, below). If the first row of the spreadsheet contains values, they will be ignored.



	1	2	3	4	5
	Day	Date	mean daily C	daily max C	daily min C
1	1.00	890101.00	-3.00	0.30	-6.10
2	3.00	890103.00	-5.60	-0.90	-9.10
3	5.00	890105.00	-4.30	2.70	-10.20
4	6.00	890106.00	-0.60	1.20	-2.10
5	10.00	890110.00	-1.70	1.40	-5.90
6	11.00	890111.00	-2.30	5.70	-6.70
7	12.00	890112.00	1.00	6.20	-4.00
8	14.00	890114.00	-4.00	-0.40	-7.20
9	15.00	890115.00	0.00	2.00	-2.50
10	17.00	890117.00	2.00	3.90	0.30

Worksheets Available

The worksheets or tables within individual spreadsheet or database files are listed here and can be chosen via the pull-down listing.

Properties

The Properties Command brings up a File Import Properties dialog window within which you can specify how GS+ should identify missing values. When importing spreadsheets and databases, other file import properties are limited to single predefined choices – e.g. spreadsheet columns are used to denote different GS+ Worksheet fields.

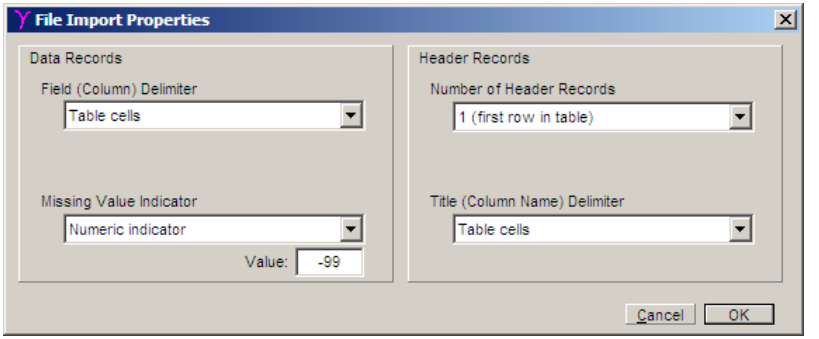
Change Column Assignments

The Change command brings up a Field or Column Assignment window from which you can assign variates (e.g. X coordinate) to columns or fields in the data file. Col-

umns can also be assigned later from the Data Worksheet window.

Input File Formats (File Import Properties)

Choosing to import files from the Import Text File or Import Spreadsheet dialogs brings up the File Import Properties window below.



Data Records
The Field (Column) Delimiter specifies how individual values within the data records are formatted. For text files values can be comma separated, tab separated, space separated (free format), character separated, or binary data. When spreadsheets are imported, columns are delimited by Table Cells only.

The **Missing Value Indicator** specifies the value or character within the file that indicates that a value is missing and that the record should be ignored during importing. A missing value that is imported becomes a blank cell in the GS$^+$ Data Worksheet. The indicator can be absent, a decimal point, a numeric value, or a character.

Header Records
The **Number of Header Records** indicates whether the first records in the file contain header records – descriptive text that should be ignored as the file is read into the Data Worksheet. For text files choose None, Varies, or Fixed Number. Header records will be read into the Data Description field of the Data Worksheet Window. Spreadsheet files are assumed to contain a single header row containing column names.

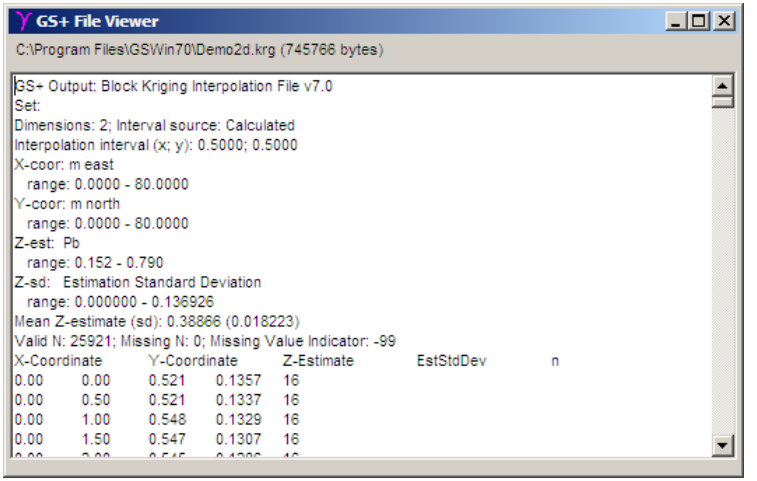
The **Column Title Separator** refers to whether column titles (variate names) appear in the second record of the file, and if so, how names are separated from one another within the record. For text files choose No Field Names, Same as for Data Records, Quotes, Brackets, Comma, Tab, Space, or Character. When spreadsheets are imported, columns titles appear in different cells of the first spreadsheet row

OK/Cancel
Press **OK** to accept the settings shown and return to the Import Text File or Import

Spreadsheet dialog. If the settings are different from the standard settings for the format originally defined, the format will change to Custom with the new settings defining the new Custom format. Press Cancel to close without importing.

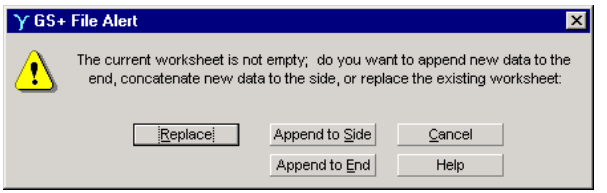
Viewing Files (File View Window)

The file view window is used to display the contents of selected files. It is available from windows in which input and output files are specified. At the top of the window is the file name and size. The contents of the window can be highlighted and copied to the Windows clipboard using the Edit – Copy menu command but cannot be changed. The example below was brought up from the View Output File command in the Interpolation window.


GS+ File Viewer
C:\Program Files\GSWin70\Demo2d.krg (745766 bytes)
GS+ Output: Block Kriging Interpolation File v7.0
Set:
Dimensions: 2; Interval source: Calculated
Interpolation interval (x; y): 0.5000; 0.5000
X-coor: m east
range: 0.0000 - 80.0000
Y-coor: m north
range: 0.0000 - 80.0000
Z-est: Pb
range: 0.152 - 0.790
Z-sd: Estimation Standard Deviation
range: 0.000000 - 0.136926
Mean Z-estimate (sd): 0.38866 (0.018223)
Valid N: 25921; Missing N: 0; Missing Value Indicator: -99
X-Coordinate Y-Coordinate Z-Estimate EstStdDev n
0.00 0.00 0.521 0.1357 16
0.00 0.50 0.521 0.1337 16
0.00 1.00 0.548 0.1329 16
0.00 1.50 0.547 0.1307 16

Data Append Dialog

When importing data into the Data Worksheet, if the worksheet already contains data then you will asked how to import the new data – whether the imported data should **Replace** the existing data, be placed in columns to the right of the existing data (**Append to Side**), or be placed at the bottom of the worksheet (**Append to End**). **Cancel** will cancel the import with no changes to the worksheet.


GS+ File Alert
The current worksheet is not empty; do you want to append new data to the end, concatenate new data to the side, or replace the existing worksheet:
Replace
Append to Side
Cancel
Append to End
Help

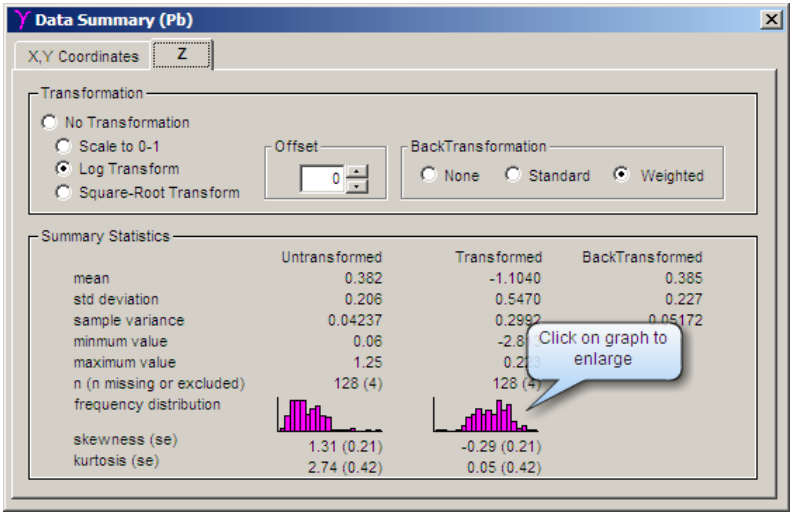
Chapter 5
Summary Statistics

Data Summary Window - Z Tab

The Data Summary window provides standard descriptive statistics for the variates defined in the Data Worksheet window. Information is provided for both the Z-variate (as below), for the coordinate variates in a separate X,Y Coordinates tab, and when a covariate is being analyzed, for the regression of Z vs. the Covariate Z2 in a separate Regression tab.

For the Z-variate it is also possible to specify a lognormal (base e) or square-root transformation in order to better normalize the variate's distribution prior to geostatistical analysis. If you do transform the variate, you may choose to have GS$^+$ report the interpolated (kriged) values either in transformed form or backtransformed to the original measurement domain. The backtransformation occurs after all analyses have been performed, and it is not applied to autocorrelation results.

Also from the Data Summary window you can access a full-window frequency distribution, a cumulative probability distribution, or a normal probability distribution by clicking on the small frequency distribution image.



Transformation

It is often helpful to apply a lognormal or a square root transformation to a Z variate in order to normalize for skewed frequency distributions. It can also be useful to scale data to a range of 0-1 if the values are extremely large. The transformation

specified is applied to every Z value in the data set prior to geostatistical analysis; the values in the data worksheet are not transformed. View the effectiveness of the transformation by viewing the Frequency Distribution or a probability distribution and the values for skewness and kurtosis in the data summary.

Offset
If your z-variates span the range of <1 to >1 (e.g. 0.3 to 20.1) and you decide to apply a lognormal or square root transform, you should make all values >1 prior to transformation by adding an offset value such as 1 (e.g. ln[z+1]). This is because of the discontinuous nature of the lognormal transformation across the <1 to >1 range.

Backtransformation
When a transformation is chosen, after analysis of the transformed data the output data are customarily (but not necessarily) back-transformed to the original data domain when reported. You may choose among three potential backtransformation choices: None, Standard, or Weighted. The standard backtransformation is simply the converse of the transformation: scaled values are rescaled to the original range, log*n*-transformed values are raised to the natural exponent *e*, and squared values are raised to the 0.5 (square root). Offset values are subtracted from the backtransformed values.

The Weighted backtransformation is a complex backtransformation that more closely approximates true population statistics than simple backtransformations. See Haan (1977) and Krige (1981) for further details.

Backtransformations are applied only to final data. These include statistics on the Data Summary screen (mean, standard deviation, etc.), and all kriging results. Individual semivariance values are not backtransformed prior to display (as noted by semivariogram axis labels).

Frequency Distribution
Click on a frequency distribution image to view an enlarged version of the frequency distribution. From the Frequency Distribution window you will also be able to view a normal probability distribution or a cumulative frequency distribution of the data (see below).

Summary Statistics
Descriptive statistics appear in the box noted. Note that n refers to the number of active data items currently in the analysis (128 in the screen above); n missing refers to the number of worksheet records that were excluded from the analysis because they contained a missing x, y, or z value or because they were excluded from the analysis by the Filter command (4 values, in the example above). If duplicate values were averaged (see the Duplicate Values dialog), then all the duplicates for a given location count as a single record, and the duplicates will not be counted as missing.

Numbers in parentheses following skewness and kurtosis are standard errors of these terms.

When analyzing data sets with a covariate (Z2) present, the following rules apply:

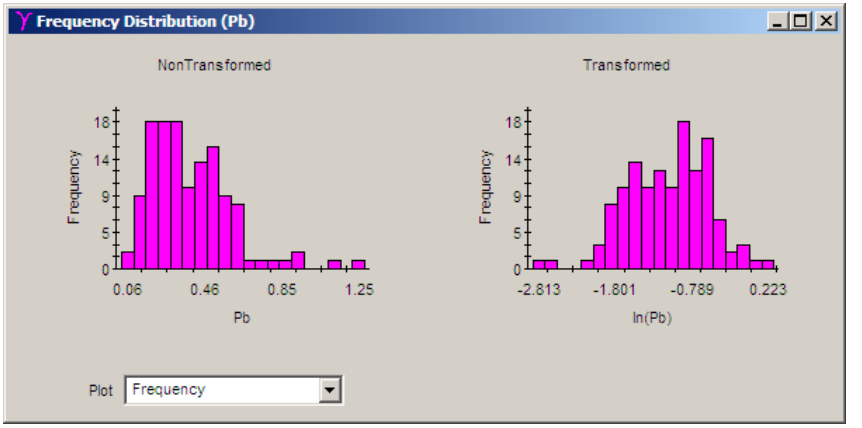
- Only records with a valid Z and a valid Z2 value will be included in the analysis of Z;
- All records with a valid Z2 value will be included in the analysis of Z2;

This is because covariance analysis expects all sample points for Z to be accompanied by a covariate Z2, which will also be sampled at places other than where Z is sampled. The summary statistics and autocorrelation analysis for Z will thus be performed only for those values of Z accompanied by a Z2. The summary statistics and autocorrelation analysis for Z2, on the other hand, will be performed for all values of Z2 regardless of whether a matching Z value is present.

Frequency Distribution Histogram

The Frequency Distribution window contains a bar graph of the frequency distribution for the Z variate. If the data are transformed, two graphs will appear, with the distribution for the transformed data to the right of the distribution for the nontransformed data (as below). If the data are not transformed, only the left-hand graph will appear. Data can be transformed in the in the Data Summary window.

The number of frequency classes (bars) can be changed using the **Edit Graph** command. In addition to a frequency distribution bar graph, you may also choose to view a Cumulative Frequency Distribution or a Normal Probability curve (described below).

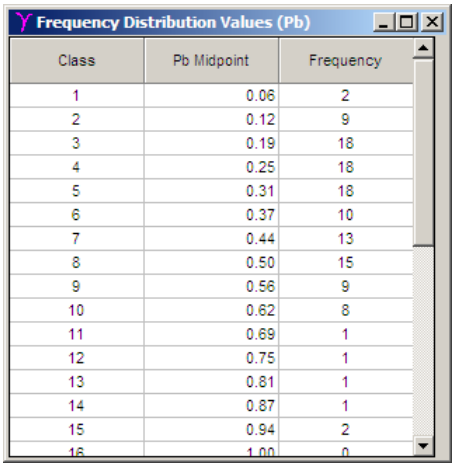


Plot
Type of plot to graph. Choose either Frequency (as displayed above), Cumulative Frequency, or Normal Probability curves (shown later).

Actions
You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS+ window, or via a right-click menu.

Frequency Distribution Values

The Frequency Listing window provides a listing of the values used to create the frequency distribution graph. The number of classes is set from Frequency Distribution window using the Edit Graph command. This is a read-only worksheet, accessed with the List values command from the right-click menu (or the Data | List graph values command from the GS+ main command menu).



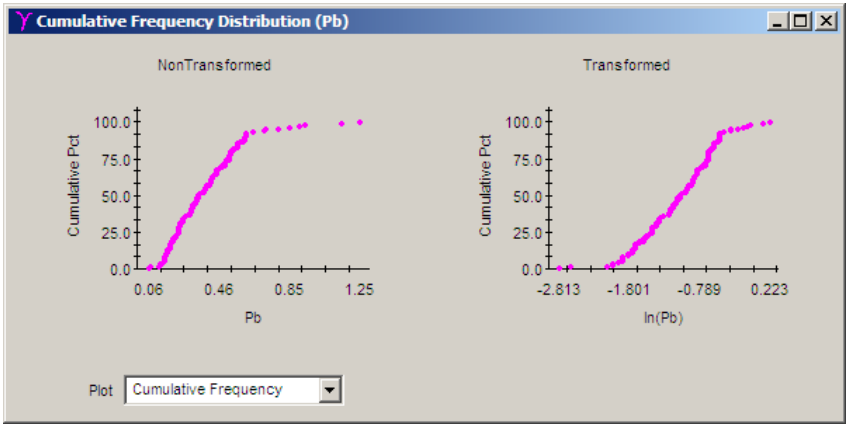
Class	Pb Midpoint	Frequency
1	0.06	2
2	0.12	9
3	0.19	18
4	0.25	18
5	0.31	18
6	0.37	10
7	0.44	13
8	0.50	15
9	0.56	9
10	0.62	8
11	0.69	1
12	0.75	1
13	0.81	1
14	0.87	1
15	0.94	2
16	1.00	0

Actions
You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS+ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Cumulative Frequency Distribution

The Cumulative Frequency Distribution window contains a graph of the cumulative frequency distribution for the Z variate. If the data are transformed, two graphs will appear, with the distribution for the transformed data to the right of the distribution for the nontransformed data (as below). If the data are not transformed, only the left-hand graph will appear. Transformations are performed in the Data Summary window.

If the data are normally distributed, the curve will describe an S-shape frequency distribution. E.G. in the graph below, transforming the data acts to normalize the distribution.



Plot

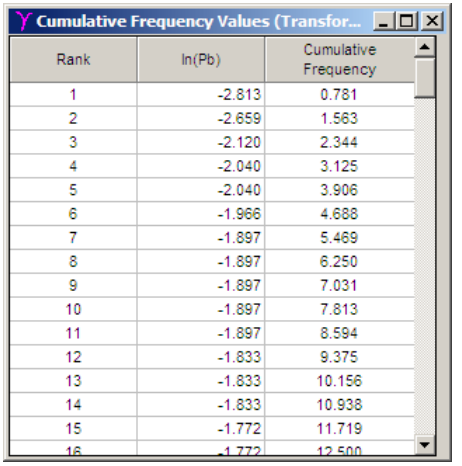
Type of plot to graph. Choose either Frequency, Cumulative Frequency (as displayed above), or Normal Probability curves (displayed below).

Actions

You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS$^+$ window, or via a right-click menu.

Cumulative Frequency Values

The Cumulative Frequency Listing window provides a listing of the values used to create the Cumulative Frequency Distribution graph. All values in the data set are ranked by value (ascending order); any given row is the frequency of the value of that row plus all preceding frequencies (expressed as percent); the final cumulative frequency value is 100%. For example, the first value in the table is the lowest value in the data set and has a cumulative frequency of 100 $\times$ (1/ n).



Cumulative Frequency Values (Transfor...

Rank	ln(Pb)	Cumulative Frequency
1	-2.813	0.781
2	-2.659	1.563
3	-2.120	2.344
4	-2.040	3.125
5	-2.040	3.906
6	-1.966	4.688
7	-1.897	5.469
8	-1.897	6.250
9	-1.897	7.031
10	-1.897	7.813
11	-1.897	8.594
12	-1.833	9.375
13	-1.833	10.156
14	-1.833	10.938
15	-1.772	11.719
16	-1.772	12.500

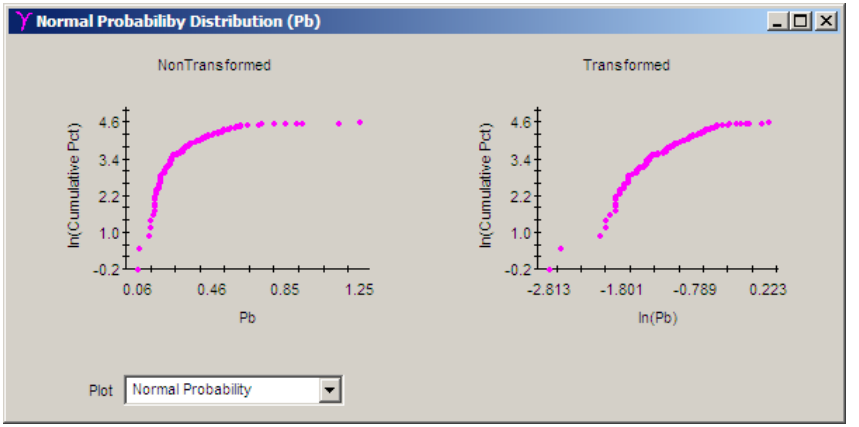
Actions

You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Normal Probability Distribution

The Normal Probability Distribution window contains a normal probability curve for the Z variate. If the data are transformed, two graphs will appear, with the distribution for the transformed data to the right of the distribution for the nontransformed data (as below). If the data are not transformed, only the left-hand graph will appear. Transformations are performed in the Data Summary window.

The normal probability curve is the same as a Cumulative Frequency distribution with the y-axis logn transformed. If the data are normally distributed, the curve will describe a straight line. E.G. in the graph below, transforming the data acts to normalize the distribution.

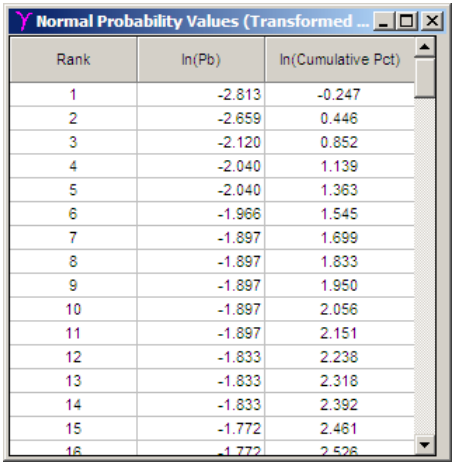


Plot
Type of plot to graph. Choose either Frequency, Cumulative Frequency, or Normal Probability (as displayed above) curves.

Actions
You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS+ window, or via a right-click menu.

Normal Probability Values

The Normal Probability Listing window provides a listing of the values used to create the Normal Probability distribution. All values in the data set are ranked by value (ascending order); any given row is the frequency of the value of that row plus all preceding frequencies (expressed as percent lognormally transformed).



Normal Probability Values (Transformed ...

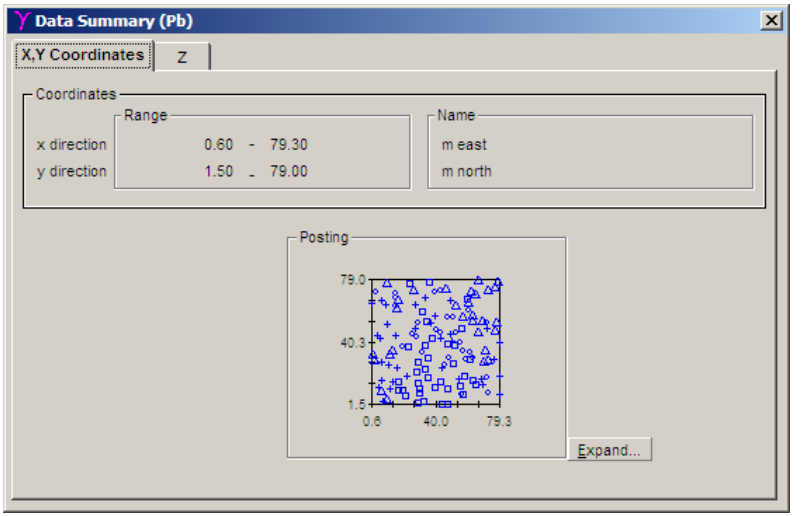
Rank	ln(Pb)	ln(Cumulative Pct)
1	-2.813	-0.247
2	-2.659	0.446
3	-2.120	0.852
4	-2.040	1.139
5	-2.040	1.363
6	-1.966	1.545
7	-1.897	1.699
8	-1.897	1.833
9	-1.897	1.950
10	-1.897	2.056
11	-1.897	2.151
12	-1.833	2.238
13	-1.833	2.318
14	-1.833	2.392
15	-1.772	2.461
16	-1.772	2.526

Actions

You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Data Summary Window - X,Y Coordinates Tab

The Data Summary provides simple descriptive statistics for the variates defined in the Data Worksheet window. Information is provided for both the X,Y Coordinates (as below) , for the Z variate in a separate Z-Variate tab (described earlier), and in a Regression tab when a covariate is assigned (described later).



Coordinates Range and Name

This is the range over which the x-direction and y-direction data vary, and the name of the variates as defined in the Data Worksheet.

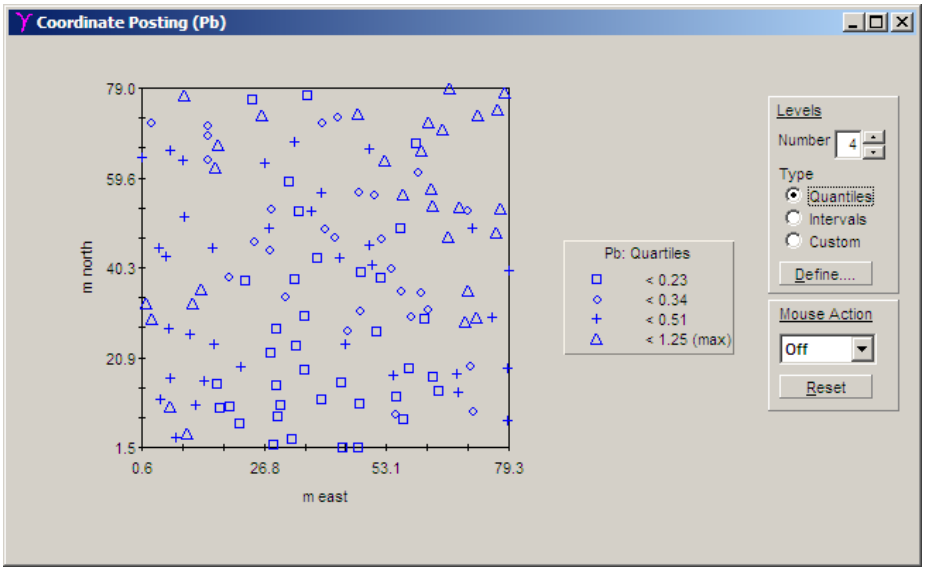
Posting

The posting is a map of the location of each x,y coordinate point within the range of X and Y coordinate values. For 1-dimensional transects the posting appears as points along a straight line. Press **Expand** or click on the map image to bring up a larger Covariate Postings (Quantile Plots) window that provides access to legends and different ways of grouping quantile data (see below). Data postings of kriged data are also available from the Map window.

Coordinate Postings (Quantile Plots)

2-d Scattergrams

The Coordinate Posting provides a map of sample locations that can be coded to show sample values at that location. Values are grouped into as many as 10 symbol classes.



Data may be grouped in one of three distributions:

Quantiles – in which Z data are placed into groups of equal frequency, i.e. into groups that have the same number of sample values. Quartile distributions are divided into 4 groups, percentiles into 100 groups, etc.

Intervals – in which Z data are placed into groups based on even intervals of the data range, without regard to how many sample values are placed in each group.

Custom – in which you can define how groups are formed.

Note that this coordinate posting is different from the posting available from the Map Window. The Map Window posting uses data saved with the interpolation file; this posting uses nonmissing data in the current Data Worksheet.

Levels
Set the **Number of levels** (corresponding to different symbols) to any value between 1 and 10. For values of 1, the legend will be suppressed.

Set the **Type of level** to correspond to how symbols are assigned to locations. **Quantiles** divide the domain into an equal number of sorted values per level. In the example above (quartiles), there are as many values greater than the median of 0.34 as there are less than 0.34.

Choose **Intervals** to divide the range into even intervals, e.g. in the example above, in which values range from 0 to 1.25, choosing Intervals would divide the data range of 0 – 1.25 into 4 levels: <0.3125, <0.625,< 0.9375, and <1.25.

Define provides the opportunity to define custom intervals as described below.

Mouse Action

Off – returns the mouse to normal operation.

Move – allows the graph to be moved within the window by clicking the left mouse button and dragging the cursor.

Scale – shrinks the graph image with the left mouse button.

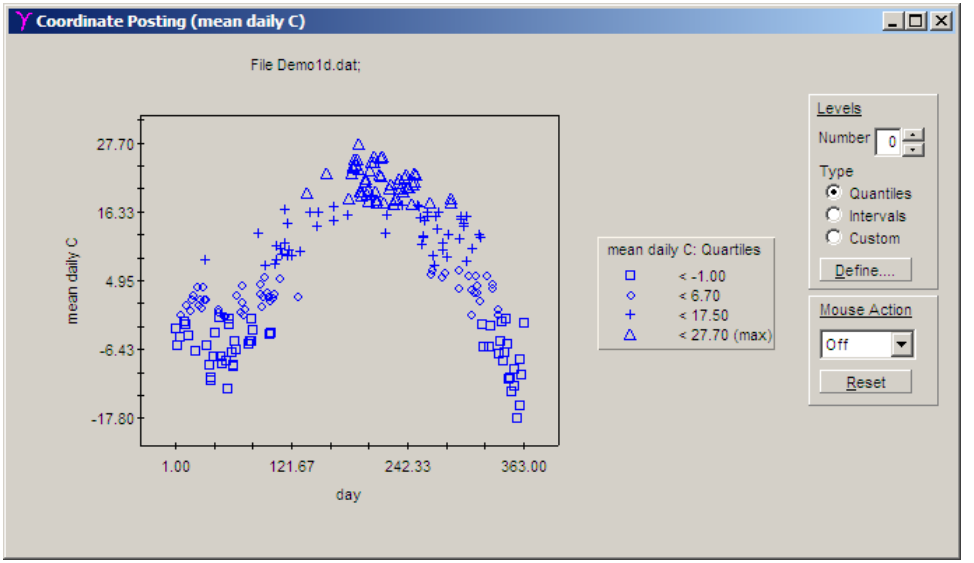
Zoom – allows you to zoom in on a particular graph area by using the left mouse button to define a rectangular zoom area. Within the zoomed area the location of the cursor is noted on the Mouse Location panel.

Reset – resets the image to the default rotation angle and scale.

Other Actions
You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS$^+$ window, or via a right-click menu.

1-d Scattergrams

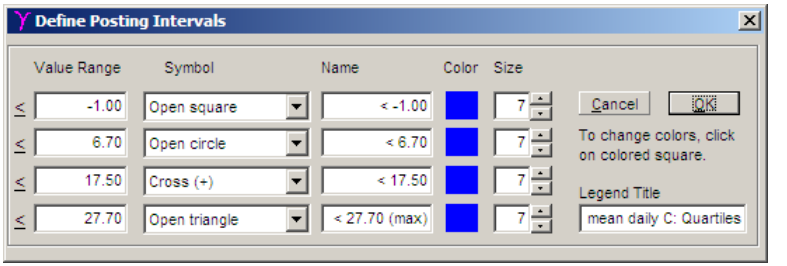
The 1-d Coordinate Posting window provides a map of coordinate locations for 1-dimensional data. The location of each data point in the active data set is marked by a symbol corresponding to its relative value. Both the number of levels displayed as well as the range for each level is user-defined. A similar window provides postings for 2-d data, as described above.



Commands in this window are the same as for 2-d Scattergrams, above.

Defining Posting Intervals

The Define Posting Intervals window allows posting intervals to be custom-defined, and provides the opportunity to set symbol size, color, shapes, and legend names. This window is accessed from the Coordinate Posting window.



Value Range
Specify the upper threshold of each value range. Changing a Percentile or Intervals value will change the Type of Level (defined in the Coordinate Posting window) to Custom. The number of values displayed is based on the number of Levels chosen in the Coordinate Posting window.

Symbol

Choose a symbol from the drop-down list for the specified interval.

Name

Choose a name to appear in the chart legend.

Color

Choose a color for the specified symbol by clicking on the colored box.

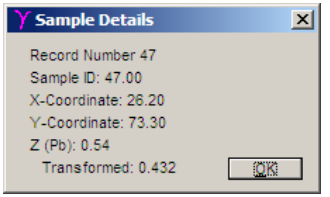
Size

Choose a size for the specified symbol

Legend title
Specify the title to appear at the top of the legends. No legend will appear if only 1 interval level is chosen.

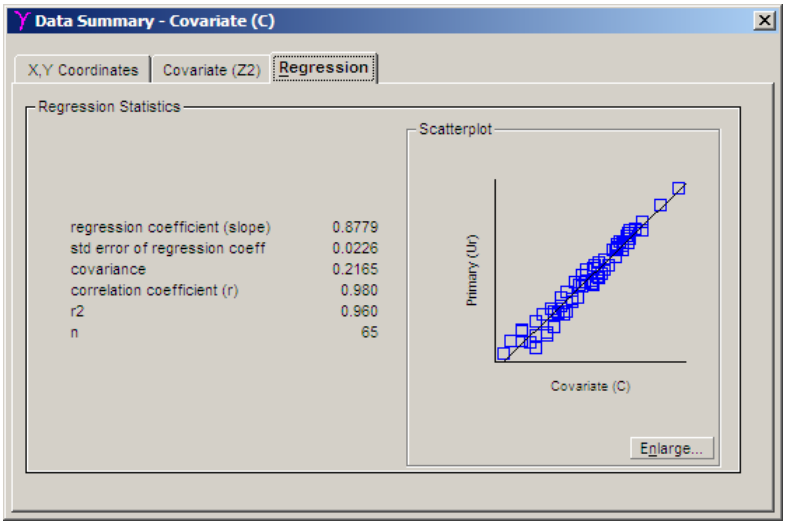
Posting Sample Details

The sample details window pops up when you click on a point in the 1-d or 2-d Co-ordinate Posting window.



Data Summary Window - Regression Analysis Tab

The Data Summary Window Regression Tab is available when a covariate is assigned in the Data Worksheet Window via the Field Assignment Dialog. Information for the Z variate or Covariate is provided in a separate Z-Variate tab, and for the coordinate variates in a separate X,Y Coordinates tab, as described earlier. The scatter plot is a plot of the cross-variate regression (see below).

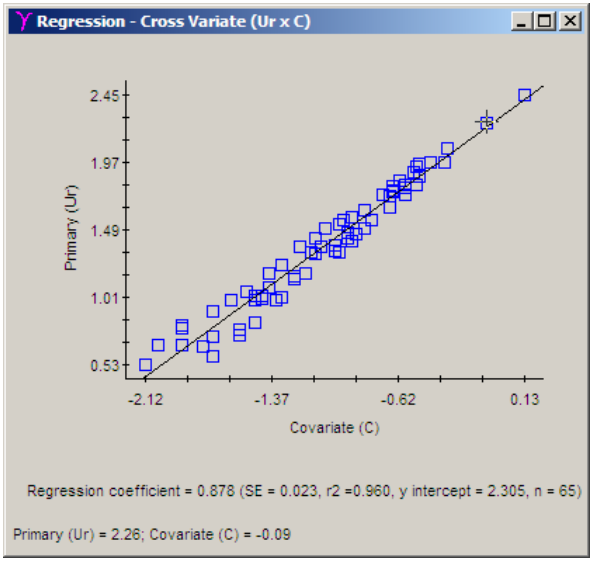


Enlarge

Press Enlarge to view a larger version of the Cross-Variate Regression scatterplot. From the Regression window you will be able to edit and print the graph as well as list graphed data values, as described below.

Cross-Variate Regression

The best test of whether a covariate is related to the primary variate – a prerequisite for cokriging – is to perform a regression analysis. In GS$^+$ the results of the regression of the primary variate Z vs. the Covariate Z2 is presented in the Data Summary Regression tab. Clicking the **Enlarge** command in that window brings up the Cross-Variate Regression window:



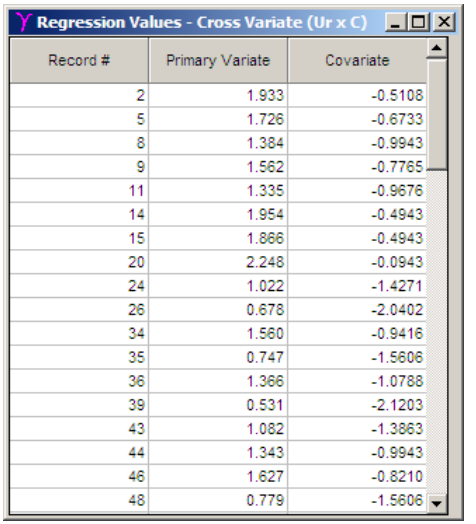
Note that when the cursor is placed over a point (the second point from the right in the graph above), the data values for that point are identified at the bottom of the window.

Actions

You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS$^+$ window, or via a right-click menu.

Cross-Variate Regression Values

The Regression Listing window provides a list of the values used to create the regression. Regression statistics appear in the Data Summary Regression tab.



Regression Values - Cross Variate (Ur x C)

Record #	Primary Variate	Covariate
2	1.933	-0.5108
5	1.726	-0.6733
8	1.384	-0.9943
9	1.562	-0.7765
11	1.335	-0.9676
14	1.954	-0.4943
15	1.866	-0.4943
20	2.248	-0.0943
24	1.022	-1.4271
26	0.678	-2.0402
34	1.560	-0.9416
35	0.747	-1.5606
36	1.366	-1.0788
39	0.531	-2.1203
43	1.082	-1.3863
44	1.343	-0.9943
46	1.627	-0.8210
48	0.779	-1.5606

Actions

You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Chapter 6
Semivariance Analysis - Overview

Semivariance is an autocorrelation statistic defined as

$$\gamma(h) = [1 / 2N(h)] \Sigma [z_i - z_{i+h}]^2$$

where

$\gamma(h)$ = Semivariance for interval distance class *h*;
z_i = measured sample value at point *i*;
z_{i+h} = measured sample value at point *i*+*h*; and
N(*h*) = total number of sample couples for the lag interval *h*.

Semivariance is evaluated in GS$^+$ by calculating γ(h) for all possible pairs of points in the data set and assigning each pair to an interval class h. For regular interval classes , GS$^+$ makes interval class assignments for any given pair of points using the following formula:

class = INT(D/DI) + 1

where

D = distance separating the pair
DI = lag class distance interval
INT = Integer function

Where the first lag is offset by ½ interval in order to better resolve values close to the origin, the formula is:

class = INT([D/DI]+0.5) + 1

This option is available by checking a box in the Preferences – Analyses dialog. It can provide better resolved variograms when there are sufficient pairs of points at shorter separation distances. The disadvantage is that if there are few pairs of points for the shortest distance classes when the first lag is not offset (a common problem), there will be even fewer pairs available with the first lag offset. If this is the case, there will be little if any improvement to the variogram.

For individually-specified lag class intervals, pairs of points are assigned to interval

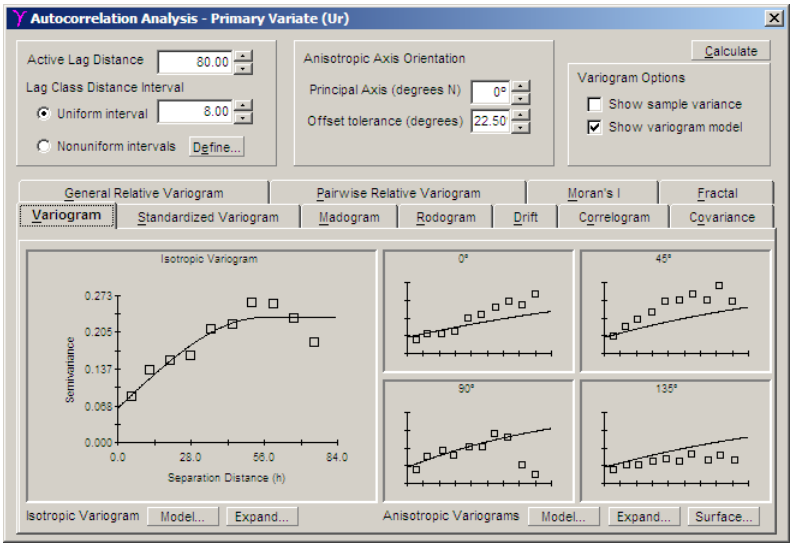
lag classes based on values in the Define Lag Class Intervals window.

GS+ calculates a semivariance statistic for each interval class; the graph of all h's vs. all semivariances for each interval class in the analysis constitutes the variogram (more properly called the semivariogram).

For cokriging, semivariance analysis must be performed for the primary variate (Z), for the covariate (Z2), and for the cross variate (Z x Z2) situation.

Autocorrelation Window

The Autocorrelation Window is where various options for variogram calculations and the end results appear; see the Semivariance Analysis summary for a definition of semivariance and formulas for lag class distance intervals.



Active Lag Distance

The Active Lag Distance specifies the range over which semivariance will be calculated. The minimum distance for this field is the minimum distance between adjacent points in the data set, while the maximum distance is the maximum distance between points. Specifying a value too large or too small will assign respectively the largest or smallest possible values to this distance.

For example, a 1200 m transect will have a maximum lag of 1200 m; specifying an Active Lag of 300 m will limit the variogram to lag intervals less than or equal to 300 m along the entire 1200 m length of the transect.

The default active lag is specified in the Analyses tab of the User Preferences dialog

window (see Chapter 2). This is not likely to be the most appropriate active lag for your data but rather will provide a starting point. Variograms typically decompose at large lag intervals because of decreasing numbers of couples per lag class as the maximum lag interval is approached.

GS^+ allows 1 million lag classes to be specified with up to 1 billion pairs per class.

Lag Class Distance Interval

The Lag Class Distance Interval defines how pairs of points will be grouped into lag classes. Each point in a variogram represents the average semivariance for a single lag class, which is a group of pairs separated by a certain Lag Class Distance Interval, sometimes called a step size. This interval can either be calculated by GS^+, in which case it will be regularly distributed across the active lag distance, or it can be manually set by the user:

- **Use individually-specified points**
  With this option you may use the **Define** command to bring up a window to Define Lag Class Intervals, i.e. to specify individual break points for the lag intervals (described below).

- **Use a regular interval**
  With this option, the value specified is the size of the interval, applied regularly across the active lag distance. E.G. an interval of 2 units with an active lag distance of 10 units will create 5 lag classes, each 2 units wide. The minimum interval allowed is the smallest distance separating any two sample point locations in the data set. The maximum interval is the greatest distance separating any two sample point locations. The default value is 10% of the active lag or, if 10% of the active lag is smaller than the minimum allowed, the minimum allowed. This default may not be appropriate for any given data set; you should try different steps for every set.

  The number of lag classes (and therefore plotted points) in a variogram is a function of values for the active lag and the active step; a 300 m active lag with a 15 m active step will have ca. 20 lag classes. Note, however, that the lag distance for a given class will be the average distance separating points within the class and not necessarily the midpoint for the class. For a 10-20 m lag class, e.g., the average lag distance may be 12.3 m rather than 15 m if more pairs of points are separated by 10-15 m intervals than by 15-20 m intervals.

  Changing the distance interval will clear results on the screen from previous analyses calculated with a different step. Results based on the new step must be re-generated with Calculate command.

Anisotropic Axis Orientation
Anisotropy refers to a direction-dependent trend in the data. Consider data collected from a two-dimensional grid on a mountain slope: elevation will be autocorrelated differently in the upslope-downslope direction than in a cross-slope direction, and thus an isotropic (all-direction) analysis may hide much of the autocorrelation that in fact is present. Anisotropic analysis allows you to see if your data have a directional component that might arise from a variety of unforeseen factors. Anisotropic analysis is irrelevant for single-dimension data such as a transect.

GS^+ evaluates geometric anisotropy, i.e. anisotropy which is expressed as variograms with different ranges in different directions. The Principal Anisotropic Axis (the Major Axis of the anisotropic model) is the direction with the longest range, i.e. the direction of major spatial continuity.

The best way to evaluate anisotropy is to view the Anisotropic Semivariance Surface (Variogram Map), and use the Azimuth function to define and then set the Principal Anisotropic Axis to the direction aligned with the lowest semivariance values (the direction of maximum spatial continuity, or major axis of the anisotropic variogram model). The map is accessed by pressing the **Surface** command at the bottom of the Semivariance Analysis window.

Principal Axis (degrees N)
The Principal Axis is the direction of maximum spatial continuity, or base axis from which the offset angles for anisotropic analyses are calculated. Offset angles are 0º, 45º, 90º, and 135º clockwise from the base axis; points aligned sufficiently close to one or another of these angles (see Offset Tolerance below) are included in the anisotropic analysis for that angle.

The axis orientation should correspond to the axis of maximum spatial continuity, i.e. the major anisotropic axis. The default axis is 0º from the north-south (y) axis. Choose the appropriate axis value by using the Surface command to bring up the Anisotropic Variogram Surface window.

Offset Tolerance (degrees)
In anisotropic analyses, the Offset Tolerance determines how closely the alignment between any two points needs to be for those points to be included in the analysis for a given offset angle. Two points will be included in the analysis for a given offset angle if the angle between them is within the offset tolerance from the offset angle.

For example, if the angle between two points is 59.3º and the offset tolerance is

15.0º, the points will be included only in the 45º angle class, which would include all angles between 30º and 60º. The default tolerance is 22.5º.

Variogram Options

Show Sample Variance
Check this option to show the sample variance for the data as a dashed line on the variogram graphs.

Show Variogram Model
Check this option to show a model for the variogram points. If the model has already been defined, either automatically or manually, the variograms will be redrawn with the model now graphed. If a model has not yet been defined, or upon executing the **Calculate** command, a best-fit model will be calculated and graphed.

To see the model parameters and to change the model, use the **Model** command at the bottom of the variogram image.

Expand
The **Expand** command brings up a separate variogram window, from which the variogram can be printed or formatted. Variance Cloud Analysis, the ability to view individual semivariance values, and the number of pairs per variogram class interval are also available from these windows.

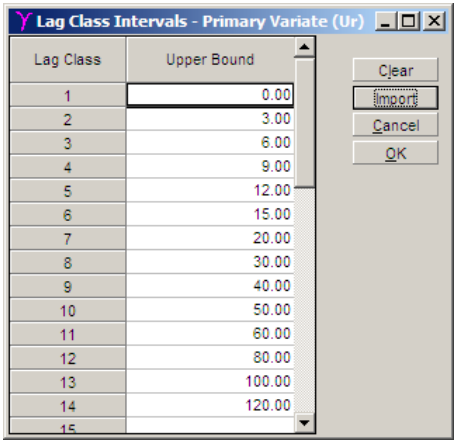
Model
The **Model** command brings up a Model Dialog window within which you may change the variogram model. The **Model** command is enabled only when the **Show Model** Variogram Option is selected.

Surface
The **Surface** command brings up the Anisotropic Variogram Surface Map window. The surface map is useful for visualizing anisotropic autocorrelation when present, as described below.

Calculate
The **Calculate** command causes the semivariogram to be calculated.

Define Irregular Lag Class Intervals

Use this dialog window to specify individual lag classes that are not necessarily regular. In the cells of the spreadsheet you can specify the upper bound of the distance interval classes desired. Zero is always the lowest bound and the Active Lag Distance is always the upper bound, regardless of values entered in this dialog. In the example below, note that classes are separated first by 3 distance units, then by 5, 10, and 20 units.



Clear
Clear the worksheet.

Import
Import a text file containing the lag class interval bounds. The default extension for lag class interval (or step) files is .stp. The format of the file to be imported is numeric-only records following a variable number of alphanumeric header records. For example:

line 1: Optional header record 1
line 2: Optional header record 2
line 3: 2.0
line 4: 4.0
line 5: 8.0
line 6: 12.0
line 7: 30.0
line 8: 100.0

This file describes 8 lag classes:

0 to <2.0,
2.0 to <4.0,
4.0 to <8.0,
8.0 to <12.0,
12.0 to <30.0,
30.0 to <100.0,
100.0 to maximum lag distance.

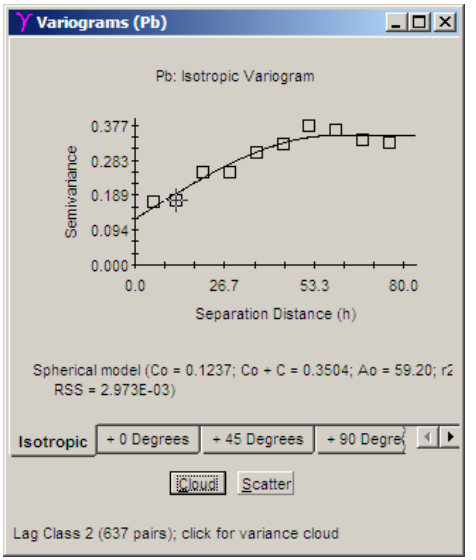
You can adjust the active lag to be any value up to and including the maximum distance separating points in the input file, i.e. the maximum lag distance.

Actions
You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Variograms Window

The Variograms window presents a full-window variogram that can be edited and printed. Different tabs hold variograms for the different anisotropic directions. Additionally, the semivariance values that were used to produce the variogram can be listed, and Variance Cloud Analysis can be performed. Variance Clouds provide a means for detecting outlier pairs of points that may be artificially skewing the variogram.

You can use the mouse cursor to identify the number of pairs in specific lag classes (reported at the bottom of the window; in the example below, the cursor was over the 2nd symbol in the variogram), and to begin variance cloud analysis. Note also that parameters for the variogram model (if present) are presented in a graph footnote. If you do not want parameters to appear as a footnote you must use the Edit Graph command to bring up the Graph Settings dialog and in the Footnote field remove the notation "(model results)".



Actions

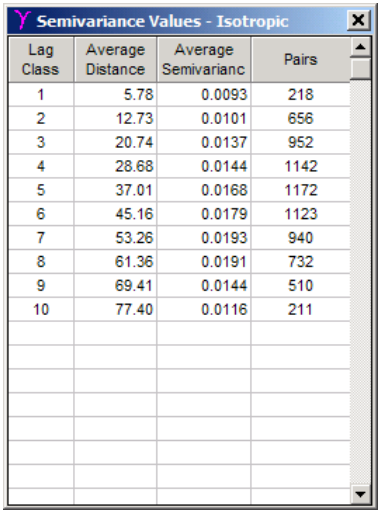
You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS+ window, or via a right-click menu.

Cloud / Scatter

Create a Variance Cloud or Scattergram graph for a particular lag interval in a separate window (described in greater detail later).

Semivariance Values

In this worksheet are listed for each lag class the average separation distance for pairs of points in that class, the average semivariance for those points, and the number of pairs of points upon which the average distance and semivariance are based. This is a read-only worksheet.



Lag Class	Average Distance	Average Semivarianc	Pairs
1	5.78	0.0093	218
2	12.73	0.0101	656
3	20.74	0.0137	952
4	28.68	0.0144	1142
5	37.01	0.0168	1172
6	45.16	0.0179	1123
7	53.26	0.0193	940
8	61.36	0.0191	732
9	69.41	0.0144	510
10	77.40	0.0116	211

Actions

You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Isotropic Variogram Models

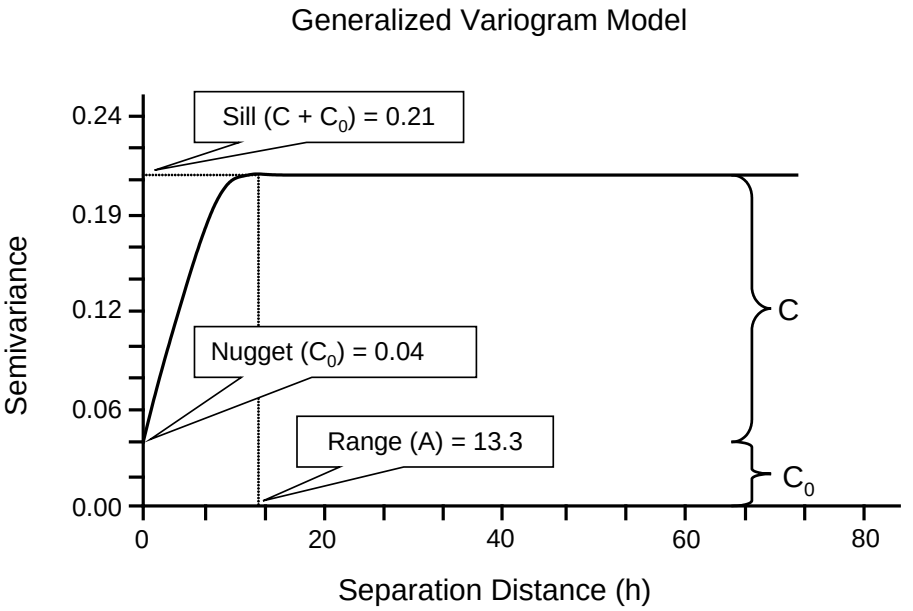
The variogram is a graph of semivariance vs. separation distance. Where autocorrelation is present, semivariance is lower at smaller separation distances (autocorrelation is greater). This typically yields a curve such as that described below, which can be modeled using three terms – a nugget variance, a sill, and a range.

GS^+ provides four types of isotropic models, each of which can be described based on three parameters:

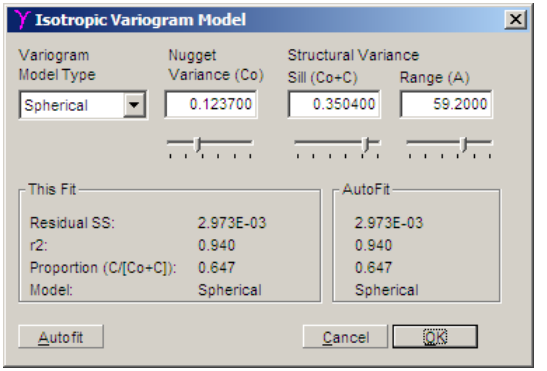
Nugget Variance or C_0 – the y-intercept of the model

Sill or C_0+C – the model asymptote

Range or A – the separation distance over which spatial dependence is apparent. Sometimes this is called the effective range in order to distinguish range (A) from a model's range parameter (A_0). In GS^+ the Range (A) is calculated from A_0 as described in the formulas for the different models, later in this chapter.



GS+ calculates default values for each parameter of the five models. You may change any of these three model parameters from the Isotropic Variogram Model dialog window:



Variogram Model Type
Choose one of four isotropic models (Linear, Spherical, Exponential, and Gaussian). As a model is chosen the variogram graphs will be updated to denote the change.

Model Terms
Any of the three model parameters for each model may be changed within the ranges allowed for individual parameters. To change a value either move the slider beneath it or type a new value directly into the text box:

Nugget Variance or C_0 – the y-intercept of the model; the nugget variance can never be greater than the sill.

Sill or C_0+C – the model asymptote; the sill can never be less than the nugget variance.

Range (A) – the separation distance over which spatial dependence is apparent. This is sometimes called Effective Range in order to distinguish range (A) from a model's range parameter (A_0). The Range A is calculated from A_0 as described in the formulas for the different models, below. The range cannot be less than 0.

Statistics – This Fit
GS+ provides three statistics to aid the interpretation of model output:

Residual Sums of Squares – provides an exact measure of how well the model fits the variogram data; the lower the reduced sums of squares, the better the model fits. When GS$^+$ autofits the model, it uses RSS to choose parameters for each of the variogram models by determining a combination of parameter values that minimizes RSS for any given model. The Residual SS displayed in the This Fit box is calculated for the currently defined model.

r^2 – provides an indication of how well the model fits the variogram data; this value is not as sensitive or robust as the Residual SS value for best-fit calculations; use RSS to judge the effect of changes in model parameters.

Proportion C/(C_0+C) -- this statistic provides a measure of the proportion of sample variance (C_0+C) that is explained by spatially structured variance C. This value will be 1.0 for a variogram with no nugget variance (where the curve passes through the origin); conversely, it will be 0 where there is no spatially dependent variation at the range specified, i.e. where there is a pure nugget effect.

Statistics – AutoFit
Statistics for the Autofit model (the model most recently calculated by GS$^+$) appear in this box. In this way you can compare your changes to model terms against that calculated automatically.

Autofit
Automatically refit the model using as starting conditions the parameters in this dialog window. Model fitting in GS$^+$ is somewhat dependent on starting conditions (assumed model parameters prior to iterations). Sometimes you can fit a better model by hand, in which case Refitting will further refine parameters to minimize RSS. To return to the original model parameters you may need to exit this Model definition window and recalculate the variogram.

OK/Cancel
Press OK to close the dialog window and apply any changes made to individual models. Press Cancel to exit the dialog window without applying changes.

Spherical Isotropic Model

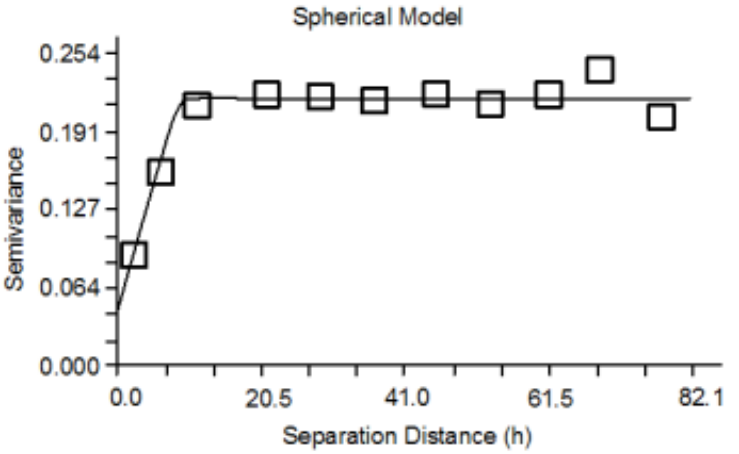
The spherical isotropic model is a modified quadratic function for which at some distance A_0 pairs of points will no longer be autocorrelated and the semivariogram reaches an asymptote. The formula used for this model is:

$$\gamma(h) = C_0 + C\,[1.5(h / A_0) - 0.5(h / A_0)^3] \quad \text{for } h \le A_0$$

$$\gamma(h) = C_0 + C \quad \text{for } h > A_0$$

where

$\gamma(h)$ = semivariance for interval distance class h,
h = the lag distance interval,
C_0 = nugget variance ≥ 0,
C = structural variance $\ge C_0$, and
A_0 = range parameter. In the case of the spherical model, the effective range $A = A_0$.



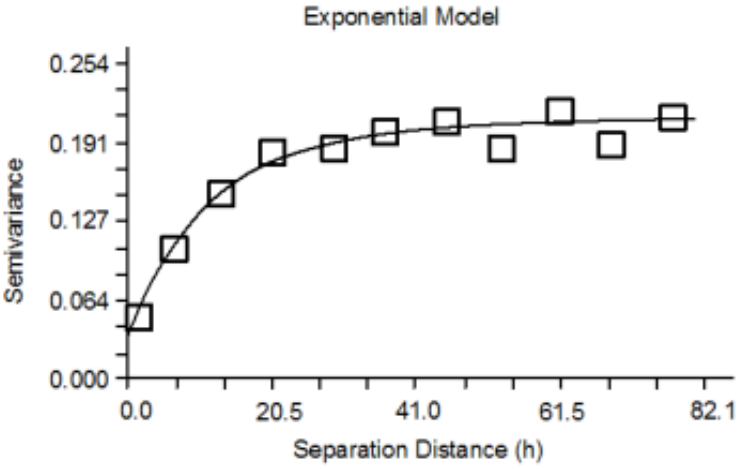
Exponential Isotropic Model

The exponential isotropic model is similar to the spherical in that it approaches the sill gradually, but different from the spherical in the rate at which the sill is approached and in the fact that the model and the sill never actually converge. The formula used for this model is:

$$\gamma(h) = C_0 + C[1 - \exp(-h / A_0)]$$

where

$\gamma(h)$ = semivariance for interval distance class h,
h = lag interval,
C_0 = nugget variance ≥ 0,
C = structural variance $\geq C_0$, and
A_0 = range parameter. In the case of the exponential model, the effective range $A = 3A_0$, which is the distance at which the sill ($C + C_0$) is within 5% of the asymptote (the sill never meets the asymptote in the exponential or Gaussian models).



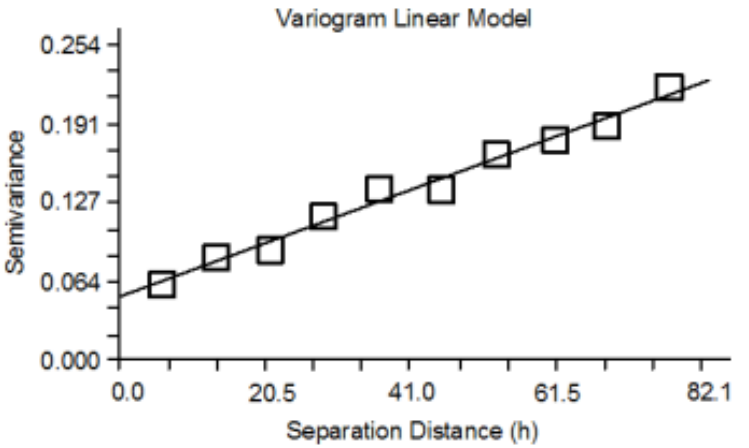
Linear Isotropic Model

The linear isotropic model describes a straight line variogram. Note that there is no sill in this model; the range A_0 is defined arbitrarily as noted below. The formula used is:

$$\gamma(h) = C_0 + [\, h(C / A_0) \,]$$

where

$\gamma(h)$ = semivariance for interval distance class h,
h = lag interval,
C_0 = nugget variance $\geq$ 0,
C = structural variance $\geq C_0$, and
A_0 = range parameter. In the case of the linear model, there is no effective range A – spatial auto-correlation occurs across the entire range sampled; in GS$^+$ model windows A is set initially to the separation distance (h) for the last lag class graphed; in the variogram below, this would be 77.4.



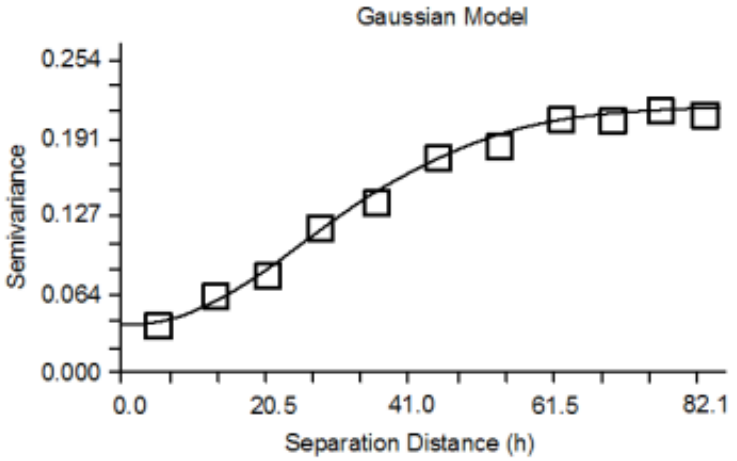
Gaussian Isotropic Model

The Gaussian or hyperbolic isotropic model is similar to the exponential model but assumes a gradual rise for the y-intercept. The formula used for this model is:

$$\gamma(h) = C_0 + C[1 - \exp(-h^2 / A_0^2)]$$

where

$\gamma(h)$ = semivariance for interval distance class h,
h = lag interval,
C_0 = nugget variance ≥ 0,
C = structural variance $\geq C_0$, and
A_0 = range parameter. In the case of the Gaussian model, the effective range $A = 3^{0.5} A_0$, which is the distance at which the sill $(C + C_0)$ is within 5% of the asymptote (the sill never meets the asymptote in the Gaussian or exponential models).



Anisotropic Variogram Models

Anisotropic variogram models are similar to those for isotropic variograms but include directional information in the range parameter. GS$^+$ calculates geometric anisotropy, in which the range changes with direction but the sill remains constant. Consequently the anisotropic model includes a directional component only for the range term below.

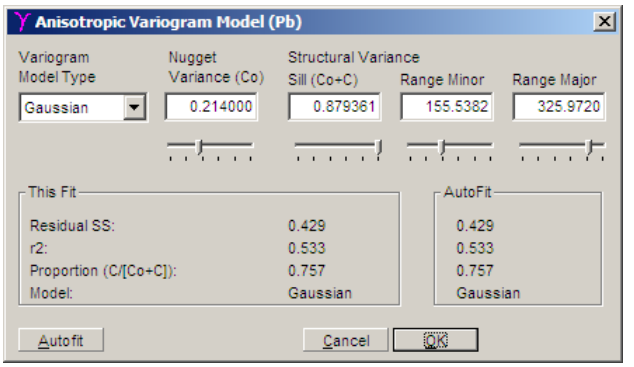
Four types of anisotropic variogram models are provided in GS$^+$: Linear, Spherical, Exponential, and Gaussian. Each model can be described using the following terms:

- Nugget Variance or C_0 – the y-intercept of the model; this value is the same for all directions.
- Sill or C_0+C – the model asymptote; this value is the same for all directions.
- Range or A – the separation distance over which spatial dependence is apparent for the direction examined. It is the sum of:
  - A_1 – the range parameter for the major axis of variation ϕ and
  - A_2 - the range parameter for the minor axis ($\phi + 90$),

    adjusted for the angle between pairs θ as noted in the formulas for individual models.

    Sometimes Range or A is called the effective range in order to distinguish range from a model's range parameters (A_1 or A_2). In GS$^+$ the Minor Ranges and Major Ranges are calculated from A_1 and A_2, respectively, as described in the formulas for the different models later in this chapter.

GS$^+$ calculates default values for each parameter of the five models. You may change any of these four model parameters from the Anisotropic Variogram Model dialog window:



Variogram Model Type
Choose one of four isotropic models (Linear, Spherical, Exponential, or Gaussian). As a model is chosen the variogram graphs will be updated to denote the change.

Model Terms
Any of the three parameters for each model may be changed within the ranges allowed for individual parameters. To change a value either move the slider beneath it or type a new value directly into the text box:

Nugget Variance or C_0 – the y-intercept of the model; the nugget variance can never be greater than the sill.

Sill or C_0+C – the model asymptote; the sill can never be less than the nugget variance.

Range Minor – the separation distance over which spatial dependence is apparent. This is sometimes called Effective Range in order to distinguish range from a model's range parameter (A_2). The minor range is calculated from A_2 as described in the formulas for the different models, below. The minor range can never be more than the major range.

Range Major – the separation distance over which spatial dependence is apparent. In some texts this is called the Effective Range in order to distinguish range from the model's major range parameter (A_1). The minor range is calculated from A_1 as described in the formulas for the different models, below. The major range can never be less than the minor range.

Statistics – This Fit
GS$^+$ provides three statistics to aid the interpretation of model output:

Residual Sums of Squares – provides an exact measure of how well the model fits the variogram data; the lower the reduced sums of squares, the better the model fits. When GS$^+$ autofits the model, it uses RSS to choose parameters for each of the variogram models by determining the combination of parameter values that minimizes RSS for any given model. The Residual SS displayed in the This Fit box is calculated for the currently defined model.

r^2 – provides an indication of how well the model fits the variogram data; this value is not as sensitive or robust as the Residual SS value for best-fit calculations; use RSS to judge the effect of changes in model parameters.

Proportion C/(Co+C) -- this statistic provides a measure of the proportion of sample variance (C_0+C) that is explained by spatially structured variance C. This value will be 1.0 for a variogram with no nugget variance (where the curve passes through the origin); conversely, it will be 0 where there is no spatially dependent variation at the range specified, i.e. where there is a pure nugget effect.

Statistics – AutoFit
Statistics for the Autofit model (the model most recently calculated by GS$^+$) appear in this box. In this way you can compare your changes to model terms against that calculated automatically.

Autofit
Automatically refit the model using as starting conditions the parameters in this dialog window. Model fitting in GS$^+$ is somewhat dependent on starting conditions (assumed model parameters prior to iterations). Sometimes you can fit a better model by hand, in which case Refitting will further refine parameters to minimize RSS. To return to the original model parameters you may need to exit this Model definition window and recalculate the variogram.

OK/Cancel
Press OK to close the dialog window and apply any changes made to individual models. Press Cancel to exit the dialog window without applying changes.

Spherical Anisotropic Model

The spherical anisotropic model is a modified quadratic function in which at some distance A_1 along the major axis and A_2 along the minor axis, pairs of points are no longer autocorrelated and the variogram reaches an asymptote. The formula used for this model is:

$$\gamma(h) = C_0 + C[1.5(h / A) - 0.5(h / A)^3] \qquad \text{for } h \leq A$$

$\gamma(h) = C_0 + C$ for $h > A$

where

$\gamma(h)$ = semivariance for interval distance class h,
h = lag interval,
C_0 = nugget variance ≥ 0,
C = structural variance $\geq C_0$, and
$A = \sqrt{\{A_1^2 [\cos^2(\theta - \phi)] + A_2^2 [\sin^2(\theta - \phi)]\}}$
A_1 = range parameter for the major axis (ϕ). In the case of the spherical model, the range (or effective range) for the major axis = A_1.
A_2 = range parameter for the minor axis ($\phi + 90$). In the case of the spherical model, the range (or effective range) for the minor axis = A_2.
ϕ = angle of maximum variation
θ = angle between pairs

Exponential Anisotropic Model

The exponential anisotropic model is similar to the spherical in that it approaches the sill gradually, but different from the spherical in the rate at which the sill is approached and in the fact that the model and the sill never actually converge. The formula used for this model is:

$\gamma(h) = C_0 + C[1 - \exp(-h / A)]$

where

$\gamma(h)$ = semivariance for interval distance class h,
h = lag interval,
C_0 = nugget variance ≥ 0,
C = structural variance $\geq C_0$, and
$A = \sqrt{\{A_1^2 [\cos^2(\theta - \phi)] + A_2^2 [\sin^2(\theta - \phi)]\}}$
A_1 = range parameter for the major axis (ϕ). In the case of the exponential model, the range (or effective range) for the major axis = $3A_1$.
A_2 = range parameter for the minor axis ($\phi + 90$). In the case of the exponential model, the range (or effective range) for the minor axis = $3A_2$.
ϕ = angle of maximum variation
θ = angle between pairs

Linear Anisotropic Model

The linear anisotropic model describes a straight-line variogram. Note that there is no sill in this model; the ranges for the major and minor axes are defined in terms of the distance interval for the last lag class in the variogram. The formula used is:

$$\gamma(h) = C_0 + h(C / A)$$

where

$\gamma(h)$ = semivariance for interval distance class h,
h = lag interval,
C_0 = nugget variance ≥ 0,
C = structural variance $\geq C_0$, and
$A = \sqrt{\{A_1^2 [\cos^2(\theta - \phi)] + A_2^2 [\sin^2(\theta - \phi)]\}}$
A_1 = range parameter for the major axis (ϕ). In the case of the linear model, there is no effective major range – spatial autocorrelation occurs across the entire range sampled; in GS$^+$ model windows A_1 is set initially to the separation distance (h) for the last lag class graphed.
A_2 = range parameter for the minor axis ($\phi + 90$). In the case of the linear model, there is no effective minor range – spatial autocorrelation occurs across the entire range sampled; in GS$^+$ model windows A_2 is set initially to the separation distance (h) for the last lag class graphed.
ϕ = angle of maximum variation
θ = angle between pairs

Gaussian Anisotropic Model

The Gaussian or hyperbola anisotropic model is similar to the exponential model but assumes a gradual rise from the y-intercept. The formula used is:

$$\gamma(h) = C_0 + C[1 - \exp(-h^2 / A^2)]$$

where

$\gamma(h)$ = semivariance for interval distance class h,
h = lag interval,
C_0 = nugget variance ≥ 0,

C = structural variance $\geq C_0$, and

$A = \sqrt{\{A_1^2 [\cos^2 (\theta - \phi)] + A_2^2 [\sin^2 (\theta - \phi)]\}}$

A_1 = range parameter for the major axis (ϕ). In the case of the Gaussian model, the range (or effective range) for the major axis = $3^{0.5} A_1$.

A_2 = range parameter for the minor axis ($\phi + 90$). In the case of the Gaussian model, the range (or effective range) for the minor axis = $3^{0.5} A_2$.

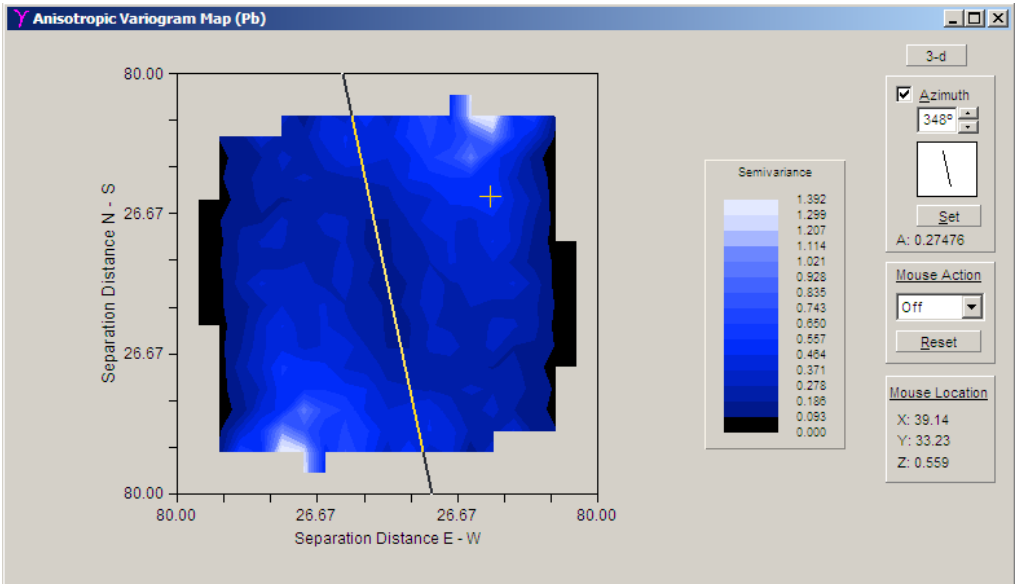
ϕ = angle of maximum variation

θ = angle between pairs

Anisotropic Semivariance Surface (2-d Variogram Map)

The Anisotropic Semivariance Surface or Variogram Map provides a visual picture of semivariance in every compass direction. This allows one to more easily find the appropriate principal axis for defining the anisotropic variogram model. A transect in any single direction (e.g. 10 degrees north) is equivalent to the variogram in that direction: the surface (z-axis) is semivariance; the x and y axes are separation distances in E-W and N-S directions, respectively. The center of the map corresponds to the origin of the variogram $\gamma(h) = 0$ for every direction. See Isaaks and Srivastava (1989; page 150) and Goovaerts (1997; page 98).

A map will appear incomplete (checkered cells, as below) when there are not at least 3 pairs of points to provide an average semivariance for a particular lag class and direction. Cells with fewer than 3 pairs will appear to be missing. Cell size is dependent on the lag class distance interval.



Actions

You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS$^+$ window, or via a right-click menu.

3-d

Change the graph projection to 3-d from 2-d or to 2-d from 3-d.

Azimuth
When the check box is marked, a transect on the map surface illustrates the azimuth angle. The angle is specified in the text box (e.g. 348º in the window above), and also displayed graphically in the area under the azimuth angle. The average semivariance for the transect is displayed as A:nnn beneath the Set command. Note that there is no difference between an angle and that angle plus 180º.

Set
Make the azimuth angle displayed in this window the Principal Anisotropic Axis in the Semivariance Analysis window. Pressing **Set** will reset the Principal Anisotropic Axis; for consistency, the angle in the Semivariance Analysis window will display a value between 0º and 180º, so that a value >180º will appear as that value less 180º (e.g. 225º will be reset to 90º in the Semivariance Analysis window).

Set will force a Rebuild of the anisotropic variograms (and models, if selected) in the semivariance window. The Principal Anisotropic Axis (the Major Axis of the anisotropic model) should be the direction of major spatial continuity (or major axis of the anisotropic variogram model), corresponding to the direction with the longest range. Correspondingly, **A** should be smaller in the major direction (lower average semivariance) and largest in the minor (90º-offset) direction.

A
The average semivariance for the azimuth transect is displayed as A:nnn beneath the Set command, where nnn is the average semivariance. **A** should be smaller in the major direction (lower average semivariance) and largest in the minor (90º-offset) direction

Mouse Action

Off – returns the mouse to normal operation.

Rotate – turns the cursor into a rotator cuff when the left mouse button is pushed, allowing the image to be rotated as desired. This option is available only for 3-d map projections.

Move – allows the graph to be moved within the window by clicking the left mouse button and dragging the cursor.

Scale – shrinks the graph image with the left mouse button.

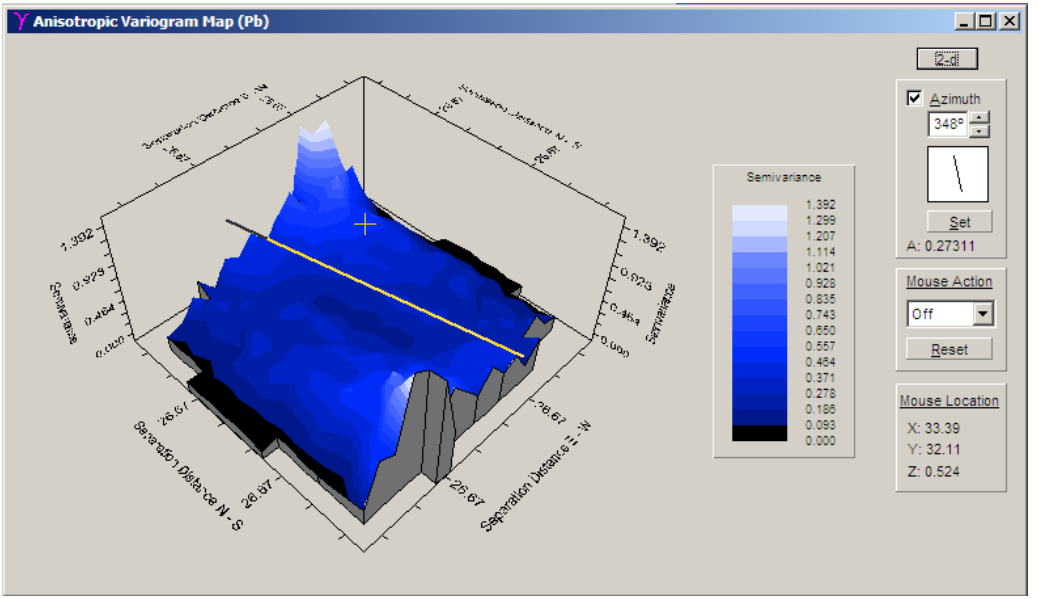
Zoom – allows you to zoom in on a particular graph area by using the left mouse button to define a rectangular zoom area. Within the zoomed area the location of the cursor is noted on the Mouse Location panel.

Reset – resets the image to the default rotation angle and scale.

Mouse Location
Provides information on the current location of the cursor when it is on the map surface. Units are map units as specified by the axes. Z is semivariance.

Anisotropic Semivariance Surface (3-d Variogram Map)

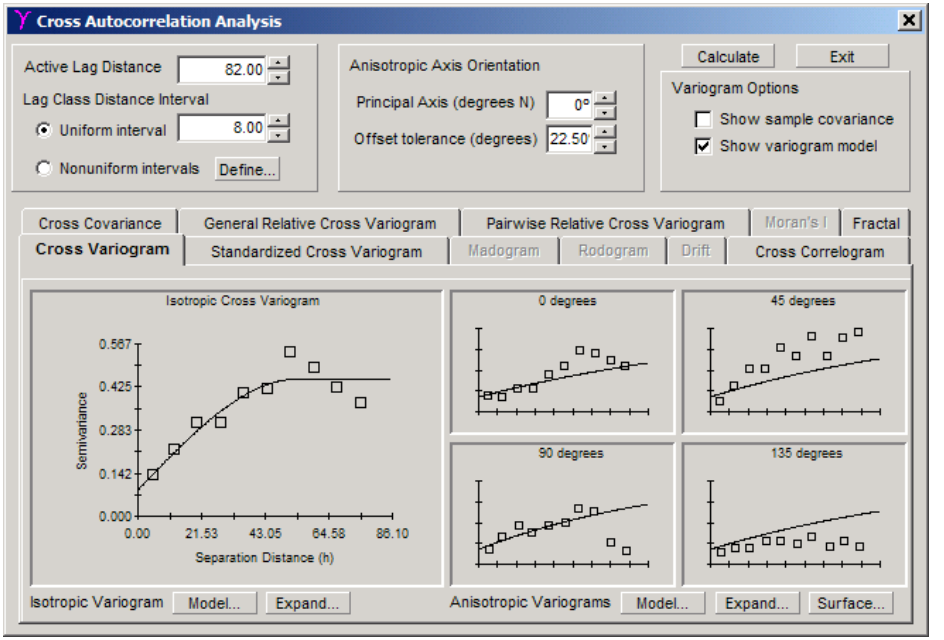
The 3-d Anisotropic Semivariance Surface or Variogram Map is a 3-d projection of the 2-d map. Use the Edit Graph command to change 3-d graph formats.



Commands for the 3-d map projection are the same as those for the 2-d projection, above.

Cross-Autocorrelation or Cross-Semivariance Analysis

The Cross-Autocorrelation Analysis Window is identical to the Autocorrelation Analysis window except that cross-semivariance analysis (Z x Z2) rather than semivariance analysis (Z or Z2) is performed. Cross-semivariance analysis must be performed prior to cokriging. For a description of the commands and options provided in this window please see the description of the Semivariance Analysis Window, above.



Commands in the Cross-Autocorrelation window are the same as in the Autocorrelation window, earlier in this chapter.

Chapter 7
Other Autocorrelation Measures

Standardized Variograms

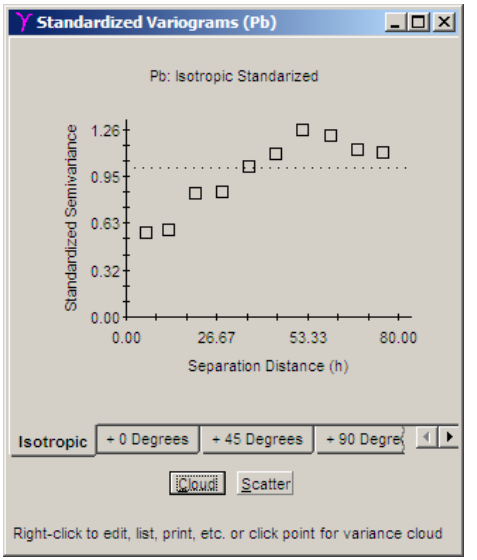
In standardized variogram analysis the variogram is computed as a proportion of sample variance:

$$\gamma(h) = \left([1 / 2N(h)] \ \Sigma [z_i - z_{i+h}]^2 \right) / s^2$$

where

$\gamma(h)$ = Semivariance for interval distance class h;
z_i = measured sample value at point i;
z_{i+h} = measured sample value at point $i+h$;
$N(h)$ = total number of sample couples for the lag interval h; and
s^2 = sample variance for all z's.

Standardized variograms (both isotropic and anisotropic) appear in the Autocorrelation Window, on the Standardized Variogram tab. Pressing the Expand button on this tab brings up the Standardized Variograms window; note the sample variance line at 1.0 along the y-axis:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Madograms

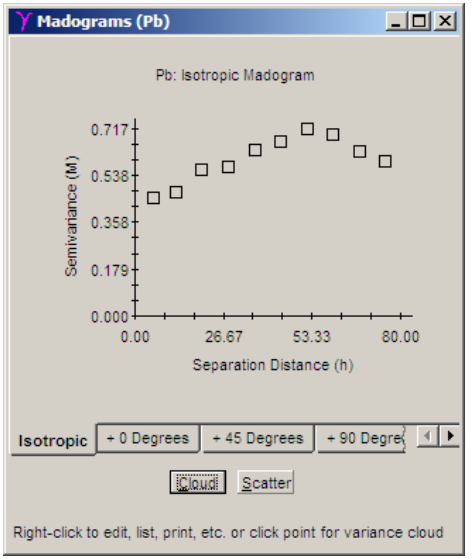
Madograms are similar to traditional variograms (see Semivariance Analysis), except that the absolute difference between z_i and z_{i+h} is used rather than the square of the difference as for traditional semivariance. The formula thus becomes

$$\gamma_M(h) = [1 / 2N(h)] \Sigma | z_i - z_{i+h} |$$

where

$\gamma_M(h)$ = semivariance(M) for interval distance class h,
z_i = measured sample value at point i,
z_{i+h} = measured sample value at point $i+h$, and
$N(h)$ = total number of sample couples for the lag interval h.

Madograms (both isotropic and anisotropic) appear in the Autocorrelation Window, on the Madogram tab. Press the Expand button on this tab to brings up the Madograms window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Rodograms

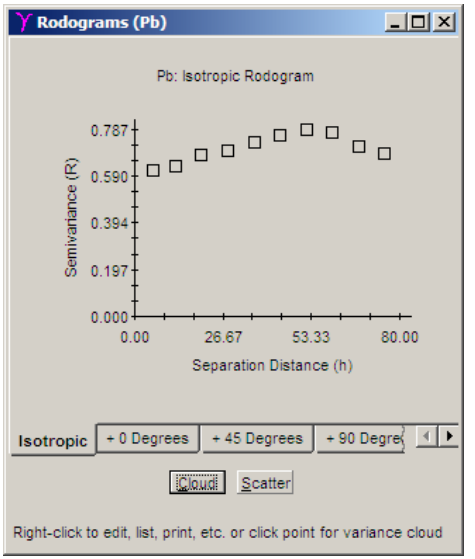
Rodograms are similar to traditional variograms (see Semivariance Analysis), except that the square root of the difference between z_i and z_{i+h} is calculated rather than the square of the difference. The formula thus becomes

$$\gamma R(h) = [1 / 2N(h)] \Sigma [|(z_i - z_{i+h})|]^{0.5}$$

where

$\gamma R(h)$ = semivariance(R) for interval distance class *h*,
z_i = measured sample value at point *i*,
z_{i+h} = measured sample value at point *i+h*, and
$N(h)$ = total number of sample couples for the lag interval *h*.

Rodograms (both isotropic and anisotropic) appear in the Autocorrelation Window, on the Rodogram tab. Press the Expand button on this to tab bring up the Rodograms window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Drift

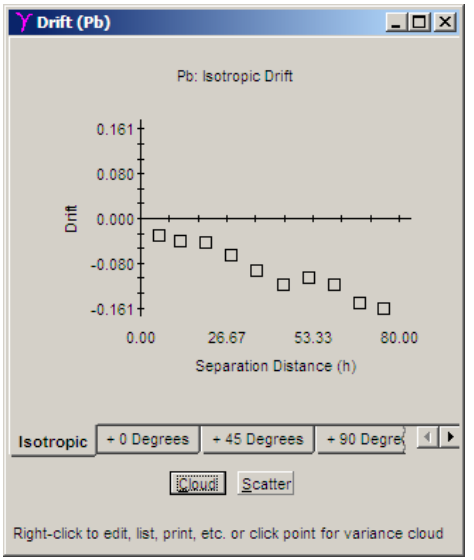
Drift (also called trend) is calculated as

$$F(h) = [1 / N(h)] \ \Sigma [(z_i - z_{i+h})]$$

where

$F(h)$ = drift for interval distance class h,
z_i = measured sample value at point i,
z_{i+h} = measured sample value at point $i+h$, and
$N(h)$ = total number of sample couples for the lag interval h.

Drift graphs (both isotropic and anisotropic) appear in the Autocorrelation Window, on the Drift tab. Press the Expand button on this tab to bring up the Drift window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Correlograms

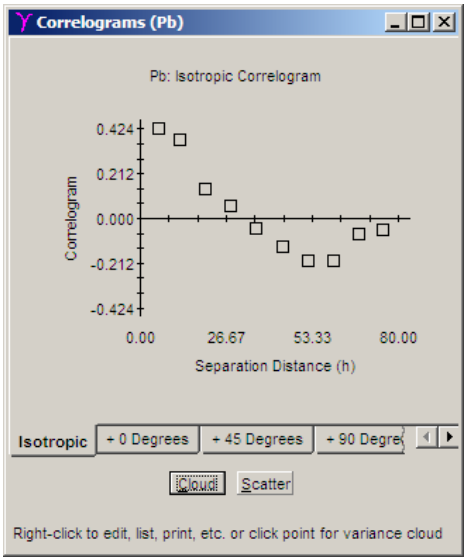
Correlograms are calculated using the formula

$$\rho(h) = C(h) / (\sigma_{-h} \quad \sigma_{+h})$$

where

$\rho(h)$ = correlation for interval distance class h,
σ_{+h} = standard deviation of all head values (z_i) for lag h, and
σ_{-h} = standard deviation of all tail values (z_{i+h}) for lag h.

Correlograms (both isotropic and anisotropic) appear in the Autocorrelation Window, on the Correlogram tab. Press the Expand button on this tab to bring up the Correlogram window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Covariance Analysis

Covariance in GS$^+$ is computed as:

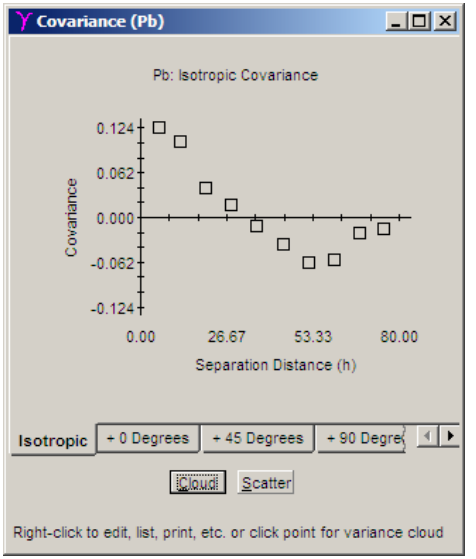
$$C(h) = [1 / N(h)] \Sigma z_i z_{i+h} - m_{-h} m_{+h}$$

where

$C(h)$ = covariance for interval distance class h,
$N(h)$ = total number of sample couples for the lag interval h,
z_i = measured sample value at point i,
z_{i+h} = measured sample value at point $i+h$,
m_{-h} = mean of all head values for lag h, or $[1/N(h)] \Sigma z_i$, and
m_{+h} = mean of all tail values for lag h, or $[1/N(h)] \Sigma z_{i+h}$.

Cross-covariance is computed in the same manner but z_i and z_{i+h} represent two different variates, Z and Z2, respectively.

The covariance graph appears in the Autocorrelation Window, on the Covariance tab. Press the Expand button to bring up the Covariance window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

General Relative Variograms

The general relative variogram is a graph of semivariance standardized by the squared mean of the data used for each distance interval h:

$$\gamma GR(h) = \gamma(h) / (m_{-h} + m_{+h})^2$$

where

$\gamma GR(h)$ = general relative semivariance for interval class *h,*

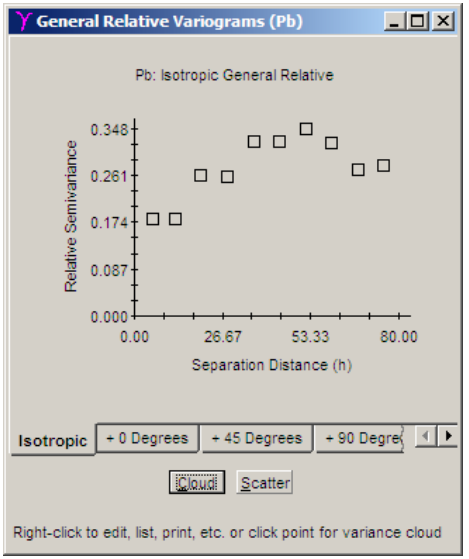
$\gamma(h)$ = semivariance for interval distance class *h*,

m_{-h} = mean of all head values for lag *h*, or $[1/N(h)] \Sigma z_i$, and

m_{+h} = mean of all tail values for lag *h*, or $[1/N(h)] \Sigma z_{i+h}$.

Relative variograms can be calculated only for data sets for which all Z values are positive (otherwise the denominator above could equal zero). If the minimum Z value in the active data is <0, then the relative variogram will not be drawn. To avoid this problem, you can first scale or otherwise transform the data in the Data Summary Window.

General relative variograms (both isotropic and anisotropic) appear in the Autocorrelation Window, on the General Relative Variogram tab. General relative variograms are not available if the minimum Z value is less than zero. Press the Expand button on this tab to bring up the General Relative Variograms window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Pairwise Relative Variograms

The pairwise relative variogram is a graph of semivariance normalized by the squared average of tail and head values:

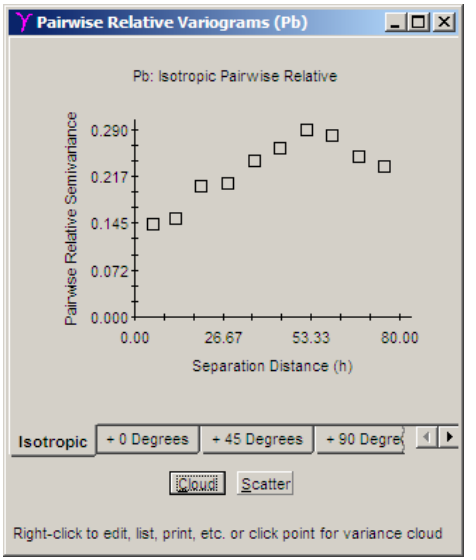
$$\gamma PR(h) = [1 / 2N(h)] \ \Sigma [(z_i - z_{i+h})^2 / [0.5(z_i + z_{i+h})]^2$$

where

$\gamma PR(h)$ = pairwise relative semivariance for interval class *h,*
$N(h)$ = total number of sample couples for the lag interval *h,*
z_i = measured sample value at point *i*, and
z_{i+h} = measured sample value at point *i+h.*

Relative variograms can be calculated only for data sets for which all Z values are positive (otherwise the denominator above could equal zero). If the minimum Z value in the active data is <0, then the relative variogram will not be drawn. To avoid this problem, you can first scale or otherwise transform the data in the Data Summary Window.

Pairwise relative variograms (both isotropic and anisotropic) appear in the Autocorrelation Window, on the Pairwise Relative Variogram tab. Pairwise relative variograms are not available if the minimum Z value is less than zero. Press the Expand button on this tab to bring up the Pairwise Relative Variograms window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Moran's I Analysis

The Moran's I statistic is a conventional measure of autocorrelation, similar in interpretation to the Pearson's Product Moment correlation statistic for independent samples in that both statistics range between -1.0 and 1.0 depending on the degree and direction of correlation. The statistic is defined as:

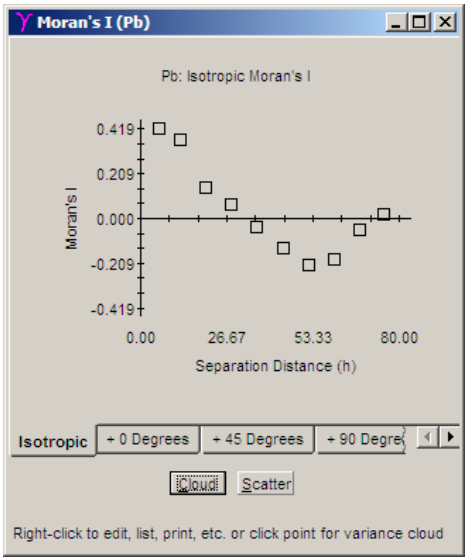
$$I(h) = N(h) \ \Sigma\Sigma z_i \ z_{i+h} \ / \Sigma z_{i+h}^2$$

where

$I(h)$ = autocorrelation for interval distance class h,
z_i = the measured sample value at point i, and
z_{i+h} = the measured sample value at point $i+h$.

Note that in this analysis all of the weights in the adjacency matrix (Sokal and Oden 1978) are set to 1, i.e. I is weighted by distance (h) between sample points rather than by simple adjacency. Calculating this statistic for a variety of lag distances yields the Moran's I autocorrelogram.

Moran's I correlograms (both isotropic and anisotropic) appear in the Autocorrelation Window, on the Moran's I tab. Press the Expand button on this tab to bring up the Morans's I window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Fractal Analysis

The Fractal Analysis window is used to set up the parameters for fractal analysis and then calculate the fractal dimension *D* for the data set defined in the Data Definition Menu. GS$^+$ calculates the fractal dimension *D* as a function of the slope of a log-log variogram plot (Burrough 1981, 1986):

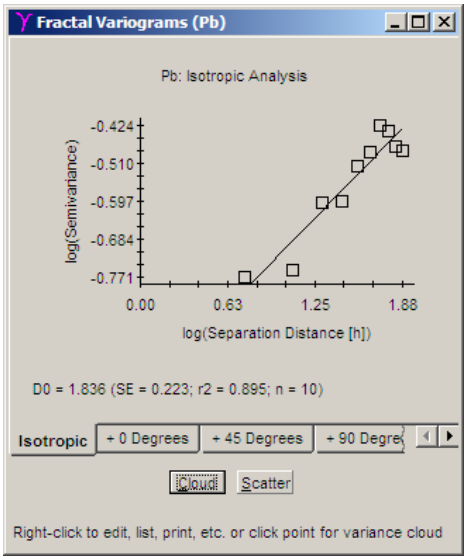
$$D = 2 - m / 2$$

where

D = the Hausdorff-Besicovitch statistic, and
m = the slope of a log-log variogram

Because *D* is based on an analysis of semivariance, it is sensitive to the same analysis parameters that affect semivariance analysis.

Fractal variograms (both isotropic and anisotropic) appear in the Autocorrelation Window, on the Fractal tab. Press the Expand button on this tab to bring up the Fractal window:



The Cloud/Scatter, right-click menu, and other commands work here as they do for the Variograms Window.

Chapter 8
Variance Clouds and h-Scattergrams

Variograms can be sensitive to outlier values in a data set, and sometimes an erratic-looking variogram can be traced to a single extreme value that inflates $\gamma(h)$ whenever they are compared with other values. One way to locate such outliers is to draw a variance cloud for each separation distance interval (or lag class) h; another way is to draw an h-Scattergram for each interval.

Variance clouds and h-Scattergrams are measures of pairwise variance in autocorrelation graphs. In a variogram each point on the graph represents the average semivariance for all pairs of points in a particular lag class. h-Scattergrams and variance clouds let you see the individual pairs of points that are used to calculate the average semivariance for that class.

Variance Clouds

The formula to calculate variance for any given pair of points at locations i and j can be reduced to the mean square difference between the points:

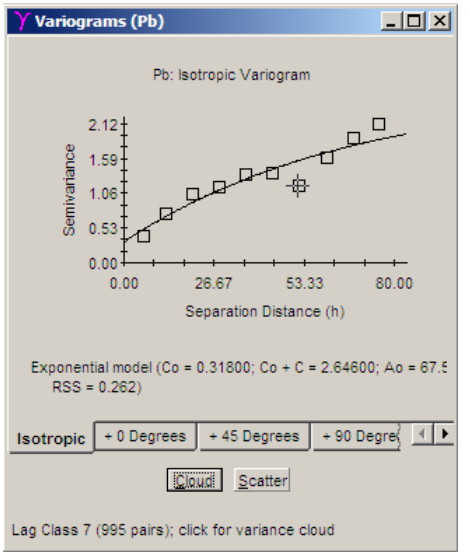
$$Var_{ij} = [\, z_i - z_j \,]^2$$

where

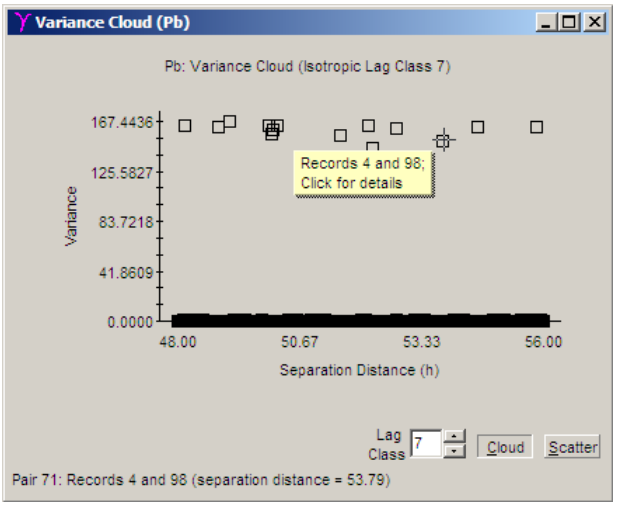
Var_{ij} = variance for pair (*i,j*),
z_i = measured sample value at point *i*, and
z_j = measured sample value at point *j*.

Any given pair(i,j) are separated by a specific distance; this distance is plotted along the x axis of the variance cloud graph. All pairs on a specific graph are in the same separation distance (lag) class.

Note that a variance cloud is specific to both direction (isotropic or a specific anisotropic direction) and to a particular lag class. In the variogram below, the cursor is on the point representing lag class 7 of the isotropic variogram, which one might suspect contains an outlier because it is so different from the other points:

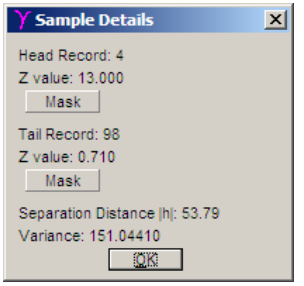


Clicking on this variogram point or the Cloud button brings up the variance cloud for lag class 7 below, and it becomes apparent that a number of pairs are very different from the others – placing the cursor over each of the outlying points reveals that all of them contain record 4 as a member of the pair. The cursor below is over the point represented by records 4 and 98 at a separation distance of 53.79, as noted at the bottom of the window:

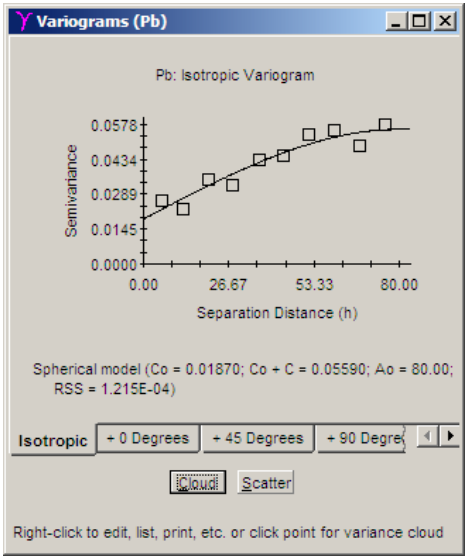


Clicking on this variance cloud point brings up the Sample Details window, which

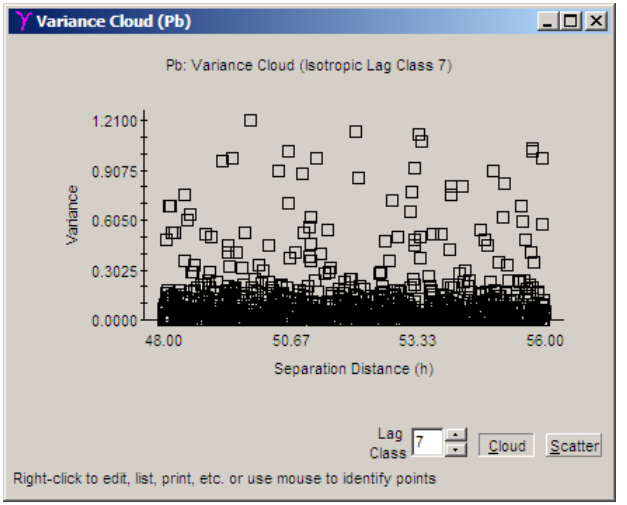
gives us the option to temporarily mask (remove from the active data set) one of the data records for this pair:



Since record 4 is a member of all of these outlier pairs, we can choose to mask record 4, which has a value of 6.0 (this can be confirmed in the Data Worksheet, although if the data are transformed the transformed value will not match the worksheet value). This gives us a more reasonable variogram:



Re-examination of the variance cloud for lag class 7 (below) reveals that the highest y-axis values are substantially lower than before (1.21 vs. 167.4) and, more importantly, all of the major outlier pairs are gone. This was accomplished by removing a single data record from the analysis:

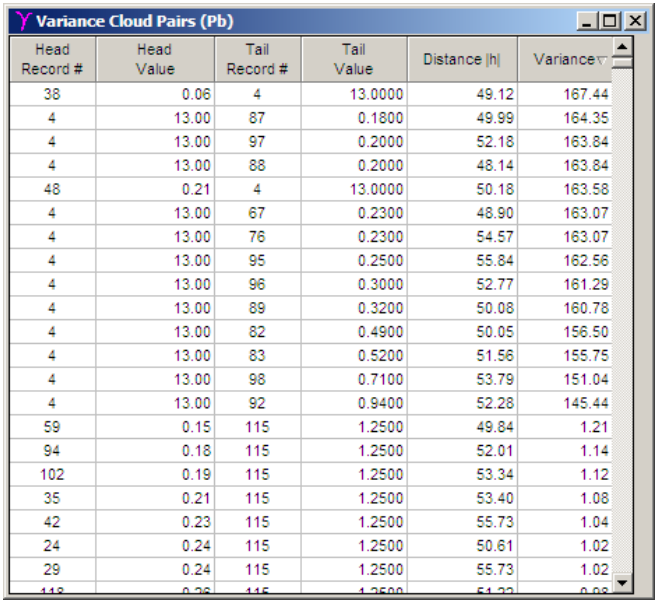


The right-click menu and other commands work here as they do for the Variograms Window.

Variance Cloud Pairs

This read-only window contains a listing of all pairs of points within a specific lag class, including for each pair the variance, separation distance, and the identity (data record number) of each member of the pair. These pairs are graphed in the Variance Cloud window.

Note that in the example below, sorted by variance (with pairs of highest variance listed first), record 4 shows up as a member of the each one of the 14 high-variance pairs. This confirms record 4 as an outlier, which was determined in the variance cloud graph, above:



Variance Cloud Pairs (Pb)

Head Record #	Head Value	Tail Record #	Tail Value	Distance |h|	Variance
38	0.06	4	13.0000	49.12	167.44
4	13.00	87	0.1800	49.99	164.35
4	13.00	97	0.2000	52.18	163.84
4	13.00	88	0.2000	48.14	163.84
48	0.21	4	13.0000	50.18	163.58
4	13.00	67	0.2300	48.90	163.07
4	13.00	76	0.2300	54.57	163.07
4	13.00	95	0.2500	55.84	162.56
4	13.00	96	0.3000	52.77	161.29
4	13.00	89	0.3200	50.08	160.78
4	13.00	82	0.4900	50.05	156.50
4	13.00	83	0.5200	51.56	155.75
4	13.00	98	0.7100	53.79	151.04
4	13.00	92	0.9400	52.28	145.44
59	0.15	115	1.2500	49.84	1.21
94	0.18	115	1.2500	52.01	1.14
102	0.19	115	1.2500	53.34	1.12
35	0.21	115	1.2500	53.40	1.08
42	0.23	115	1.2500	55.73	1.04
24	0.24	115	1.2500	50.61	1.02
29	0.24	115	1.2500	55.73	1.02

Actions

You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

h-Scattergrams

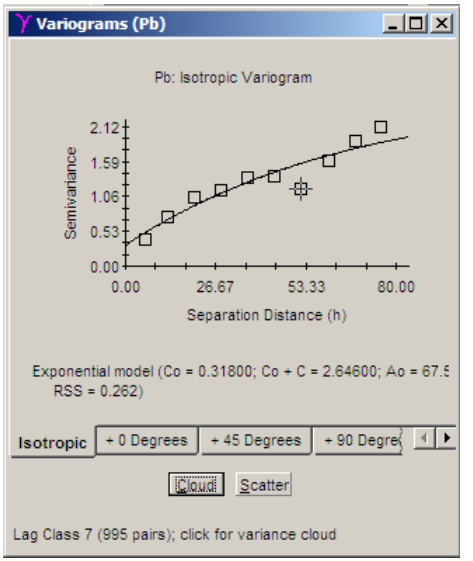
In an h-Scattergram each head value z_i is plotted against each tail value z_j for a given lag class h. The formula to calculate this difference for any given pair of points at locations i and j is

$$D_{ij} = z_i - z_j$$

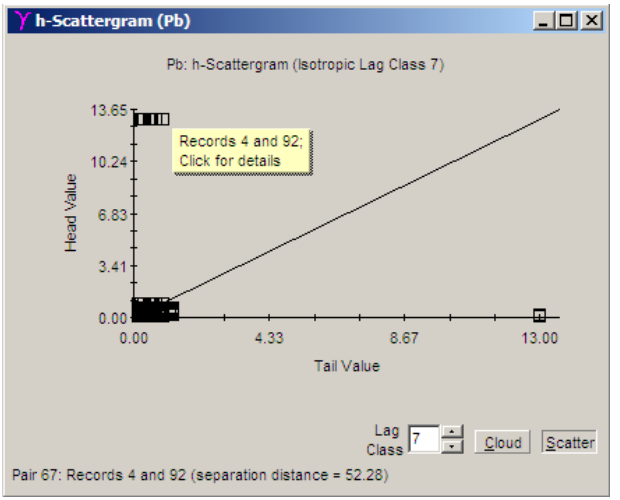
where

D_{ij} = difference for pair (*i,j*),
z_i = measured sample value at point *I*, and
z_j = measured sample value at point *j*.

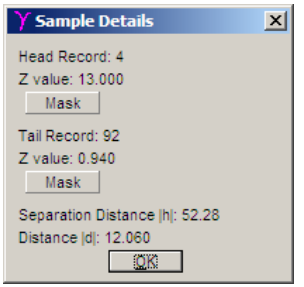
Note that an h-Scattergram is specific to both direction (isotropic or a specific anisotropic direction) and to a particular lag class. In the variogram below, displayed earlier, the cursor is on the point representing lag class 7 of the isotropic variogram which looks odd:



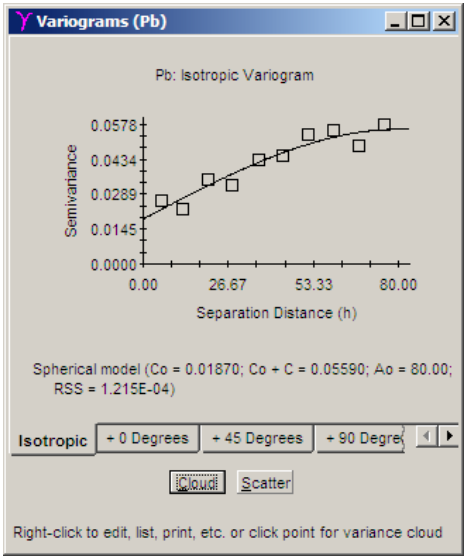
Clicking on this variogram point or pressing the Scatter button brings up the h-Scattergram for lag class 7 below, and it becomes apparent that a number of pairs are very different from the others. Placing the cursor over each point reveals that all of the points farthest from the 45-degree line contain record 4 as a member of the pair (the cursor below is over the point represented by records 4 and 92 with a separation distance of 52.28):



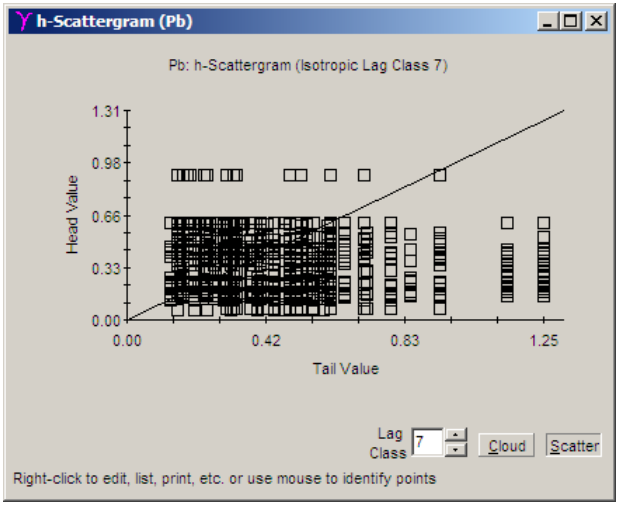
Clicking on this h-Scattergram point brings up the Sample Details window, which gives us the option to temporarily mask (remove from the active data set) one of the data records for this pair:



Since record 4 is a member of all of these outlier pairs, we can choose to mask record 4, which has a value of 6.0 (this can be confirmed in the Data Worksheet, although if the data are transformed the transformed value will not match the worksheet value). This gives us a more reasonable variogram:



Re-examination of the variance cloud for lag class 7 (below) reveals that the highest y-axis values are substantially lower than before (1.31 vs. 13.65) and, more importantly, all of the major outlier pairs are gone. This was accomplished by removing a single data record from the analysis:

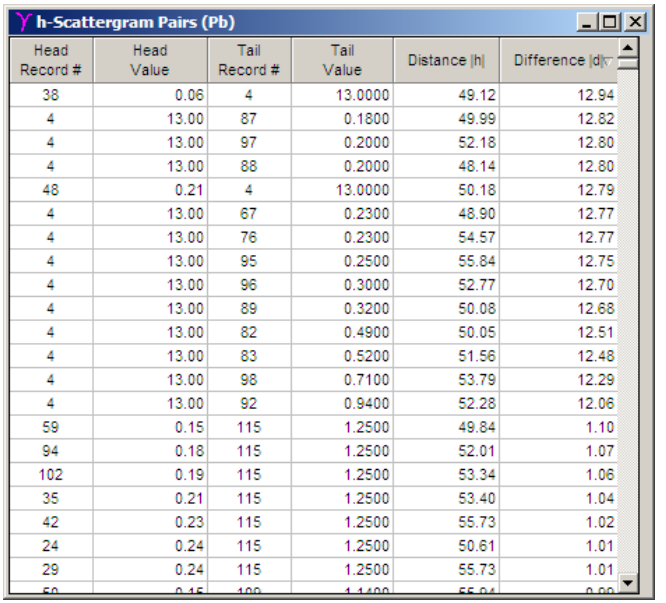


The right-click menu and other commands work here as they do for the Variograms Window.

h-Scattergram Pairs

This read-only window contains a listing of all pairs of points within a specific lag class, including for each pair the difference, separation distance, and the identity (data record number) of each member of the pair. These pairs are graphed in the h-Scattergram window.

Note that in the example below, sorted by difference (with pairs of greatest difference listed first), record 4 shows up as a member of 14 high-difference pairs. This confirms the record as an outlier, as was found in the h-Scattergram above:



h-Scattergram Pairs (Pb)

Head Record #	Head Value	Tail Record #	Tail Value	Distance |h|	Difference |d|
38	0.06	4	13.0000	49.12	12.94
4	13.00	87	0.1800	49.99	12.82
4	13.00	97	0.2000	52.18	12.80
4	13.00	88	0.2000	48.14	12.80
48	0.21	4	13.0000	50.18	12.79
4	13.00	67	0.2300	48.90	12.77
4	13.00	76	0.2300	54.57	12.77
4	13.00	95	0.2500	55.84	12.75
4	13.00	96	0.3000	52.77	12.70
4	13.00	89	0.3200	50.08	12.68
4	13.00	82	0.4900	50.05	12.51
4	13.00	83	0.5200	51.56	12.48
4	13.00	98	0.7100	53.79	12.29
4	13.00	92	0.9400	52.28	12.06
59	0.15	115	1.2500	49.84	1.10
94	0.18	115	1.2500	52.01	1.07
102	0.19	115	1.2500	53.34	1.06
35	0.21	115	1.2500	53.40	1.04
42	0.23	115	1.2500	55.73	1.02
24	0.24	115	1.2500	50.61	1.01
29	0.24	115	1.2500	55.73	1.01

Actions

You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Chapter 9
Interpolation Basics

Interpolation is the estimation of values in an area for points not actually sampled. There are many different interpolation techniques, ranging from simple linear techniques that average the values of nearby sampled points, to more complex techniques like kriging that use in the average weights based on distance to nearby sample points and the degree of autocorrelation for those distances.

GS$^+$ provides four broad types of interpolation. All are nearest-neighbor techniques in which values at locations close to the interpolation point are used to estimate the interpolation point value. They differ in the way that nearby locations are weighted and interpolations calculated. The four techniques are:

Kriging, in which interpolation estimates are based on values at neighboring locations plus knowledge about the underlying spatial relationships in a data set. Semivariograms provide knowledge about the underlying relationships. The estimated value at a given location is a weighted moving average of best estimates calculated to minimize local area variance. There are a number of different types of kriging, described below.

Cokriging, in which kriging interpolations include a covariate that is related to the primary variate and is measured at more locations than the primary variate. In cokriging there is a variogram for the primary variate, for the covariate, and for the cross-variate.

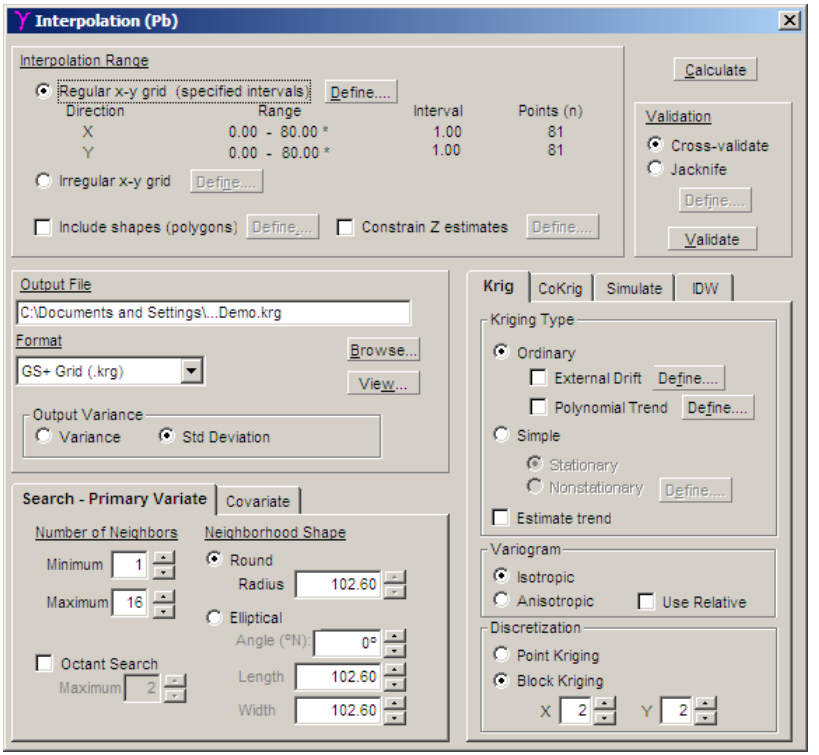
Conditional Simulation, in which interpolations are based on a form of stochastic simulation in which data values are honored at their locations. This means that local details are not obscured by smoothing as they are in kriging.

Inverse Distance Weighting (IDW) and **Normal Distance Weighting (NDW)**, in which interpolation estimates are made based on values at nearby locations weighted only by distance from the interpolation location. Neither IDW nor NDW make assumptions about spatial relationships except the basic assumption that nearby points ought to be more closely related than distant points to the value at the interpolate location. IDW applies stronger weights to nearby points than does NDW.

All of these techniques are described in greater detail on the following pages.

The Interpolation Window

The Interpolation Window provides access to Kriging, Cokriging, Conditional Simulation, and Inverse Distance Weighting techniques for estimating values for points not sampled. Interpolation produces an output file that is used by GS+ for mapping. The output file can also be read by other mapping programs.



Interpolation Range

Defines where to place interpolation estimates – in a regularly spaced grid across a rectangular area or at user-specified locations, in either case with or without masks that can define areas to include or exclude.

Regular x-y grid (specified intervals)

A grid is defined by a rectangle that has an X-direction length, a Y-direction length, and for each direction, intervals between the grid intersections. Interpo-

lation locations are at every grid intersection.

The default range is defined by the minimum and maximum X-coordinate and Y-coordinate values, and an interval based on a certain number of points in each direction. For 1-dimensional data sets, there is no y direction.

The grid can be changed with the **Define** command, which will display an Interpolation Grid dialog window.

Irregular x-y grid (specified points)

An irregular grid is composed solely of interpolation locations specified by the user. Press Define to bring up an Interpolate Worksheet within which locations can be defined or imported from an external text file.

Include shapes (polygons)

Specific shapes can be interpolated or excluded from being interpolated by defining inclusive or exclusive polygons prior to kriging. Press Define to bring up a Polygon Outlines Worksheet within which to define polygons. See Define Polygon Outlines later in this chapter.

- In inclusive polygons, the area within the polygon is interpolated (or mapped);
- In exclusive polygons, the area within the polygon is not interpolated (or mapped).

Constrain Z estimates

The estimated Z values (interpolates) can be constrained to a specific range. For example, if values less than zero are inappropriate, you can specify that Z estimates less than zero be reported as zero. Press Define to bring up a dialog box to specify Z Estimate Boundaries, described later in this chapter.

Output File Name
Press **Select** to select an existing or new file to which kriging interpolation estimates will be written. To examine the contents of an existing file press **View**.

Output Format
The format with which GS$^+$ will write estimates to the file can be one of several types.

GS+ format (.krg) – in this format a header area defines the interpolation grid, variate names, and other information about the file needed to initiate mapping later, and the data records include for each X and Y Coordinate location that is kriged the interpolation or Z-estimate, the standard deviation of the Z-estimate, and the number of neighbors that were used to make the estimate. See detailed example below.

Surfer® Grid format (.grd) – in this format a short header area defines information needed for mapping, and the data is written as a continuous stream of Z-estimates beginning from the NW corner of the interpolation grid. The standard deviation of the estimate and the number of neighbors used for interpolation are NOT included in this format. This format is compatible with Golden Software's Surfer mapping program. Note that this format is not the same as the Surfer XYZ Input file format. See detailed example below.

ArcView® Format (.asc) – this is similar to the Surfer format but the header area is formatted differently and the Z-estimates are written in a pattern that begins from the NW corner of the interpolation grid. The standard deviation of the estimate and the number of neighbors used for interpolation are not included in this format. **Also for this format, the x and y interpolation intervals must be the same** (you can set them to be the same from the Interpolation Grid dialog window). This format is compatible with ESRI's ArcInfo® Geographic Information System. See detailed example below.

GSLib Format (.out) – this format is similar to the GeoEas input format. A long first record contains coordinate interval information that is read by GS+ when mapping GSLib files. The second record of the file specifies three header records named "estimate," "estimation variance" (or "standard deviation"), and "neighbors." Records that follow are in the same order as for Surfer®, but there are three fields per record (estimate, estimation error, and number of neighbors) rather than just one field (the Z-estimate). See detailed example below.

Output Variance
You may choose to report estimation error as estimation variance or as estimation standard deviation.

Search Neighborhood
GS+ interpolates values for a specific location using nearest neighbor values weighted by distance. Only a certain number of near neighbors are used to calculate the interpolation estimate, and neighbors can be required to be within a particular geographic area around the location being estimated.

The default value of 16 nearest neighbors is usually sufficient, and also by default no

restrictions are placed on the neighborhood radius (in Kriging, neighbors outside of the variogram range are weighted identically and, if significant structural dependence is present, weighted minimally). Specifying more than 16 neighbors can slow interpolation substantially, although up to 64 neighbors are allowed. The geographic area to be searched can be either round or elliptical; if elliptical, the angle of the ellipsoid in addition to its width and length must be specified – the default angle is the same as the anisotropic axis orientation used for semivariance analysis.

Additionally, you may specify an octant search, in which the geographic neighborhood is divided into 8ths, and only a limited number of neighbors within the octant are used for interpolation. This is useful when there are many more neighbors on one side of an estimation location than on another, and there is a danger of using a disproportionate number of neighbors from that side. An octant search limits the number of neighbors that can be used from any given octant.

When cokriging, you may specify a separate set of search criteria for the covariate (Z2). These criteria are set on the Covariate tab of the Search box.

Interpolation Type – Krig, Cokrig, Simulate, or IDW
You may choose to interpolate using either Kriging, Cokriging, Conditional Simulation, or IDW (Inverse Distance Weighting) by choosing one or the other tabs on the right hand side of the window; the options on each tab are described in the next section.

Calculate
Perform interpolation analysis. During analysis values are written to the output file in the specified format. You can stop analysis by pressing the <Escape> key or the Cancel button on the progress bar.

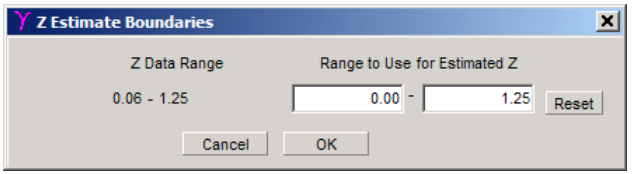
Validation
Two types of validation are provided by GS$^+$: **Cross-validation** and **Jackknife** analysis. In cross-validation analysis each measured point in the spatial domain is individually removed from the domain and its value estimated as though it were never there. Then the point is replaced and the next point is removed and estimated, and so on. In this way a graph can be constructed of estimated vs. actual values for each sample location in the domain.

In Jackknife analysis, estimates are compared against measured values for a set of locations different from those used as input data. The jackknife data are specified in a worksheet that appears when the “Define” command is pressed.

Each of these validation analyses are described more thoroughly below.

Define Z Estimate Boundaries

The Z Estimate Boundaries window is used to constrain estimated Z values (interpolates) to a specific range. It is accessed from the Interpolation Window.



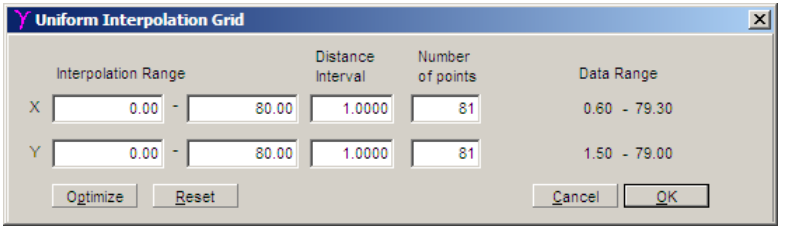
Range to Use for Estimated Z
Enter the lowest permissible value for estimated Z in the left-hand box, and the highest permissible value in the right-hand box. In this example, if the interpolation system were to estimate a value of -0.45 for a particular point, the value will be reported as 0.

Reset
Reset values to the lowest and highest data values, respectively.

OK/Cancel
Press OK to save changes and close the window; press Cancel to close the window without saving changes.

Defining a Regular Interpolation Grid

With the Regular Interpolation Grid dialog window you may define the region to be kriged or interpolated as well as the intensity at which the interpolation is to take place. This method for specifying interpolate locations is appropriate for interpolations at regular intervals across an area; if the interpolations are to be performed at odd locations, use the Irregular x-y Interpolation Grid. If the outline of the area is a complex polygon you should still use this dialog (Regular Interpolation grid) but then specify the shape of the polygon from the Interpolation window (see Polygons, below). Note that for 1-dimensional data sets (e.g. transects or time series) only the X direction is displayed in the dialog window.



Interpolation Range
Specify the beginning and ending values for the region to be interpolated. The region may exceed the data range, which is noted on the right side of the dialog window. Irregularly-shaped areas within the bounds of the interpolation range may be interpolated by specifying polygon masks from the Krig window.

The number of decimal places used to define the range is limited to the number of decimal places specified for the given coordinate field in the Data Worksheet Window. Changing the number of decimal places for a coordinate in the Data Worksheet Window changes the number of decimal places reported here. Changing the range will change the Number of points value.

Note that interpolation will begin at the minimum X and Y values, but may not reach the maximum values if the total distance (e.g. the maximum X value less the minimum X value) is not evenly divisible by the distance interval. For example, interpolation of an X Range of 0 – 105 with a distance interval of 10 will stop interpolating at X = 100, prior to exceeding the maximum range of 105.

Data Range
This is the range covered by the actual data set; these are read-only values, they cannot be changed by the user from this dialog window.

Distance Interval
Specify the distance interval between locations within the interpolation range. A distance interval of 2.0 over a range of 0 to 10.0 means that interpolations will be made at points 0.0, 2.0, 4.0, 6.0, 8.0, and 10.0. Changing the distance interval will change the value for number of points.

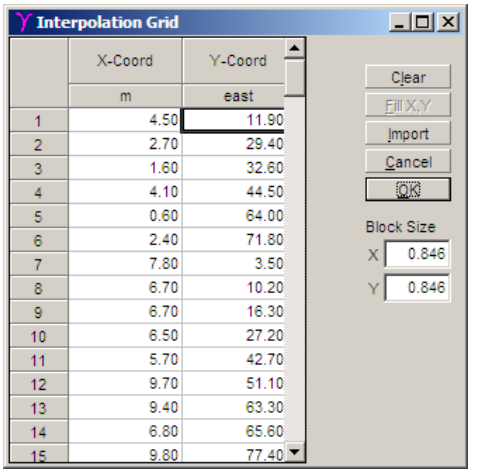
Number of Points
For any given range to be interpolated, the number of points will be the range divided by the distance interval plus 1. For a range of 0 to 10.0 with a distance interval of 2.0, the number of points will be [(10.0 – 0)/2.0] +1 = 6 points at locations 0.0, 2.0, 4.0, 6.0, 8.0, and 10.0.

Optimize
The Optimize command sets the number of points to 61 for both the X and Y direc-

tions (101 points for a single-dimension transect) and calculates the appropriate Distance Interval. The Interpolation Range is not affected.

Defining an Irregular Interpolation Grid

Specify in this worksheet the locations of individual points to be kriged. This method of specifying interpolate locations is appropriate when locations are not regularly spaced across the interpolation area. If interpolate locations are regularly spaced, regardless of whether the area is rectangular or a complex polygon, it is usually more efficient to use a Regular Interpolation Grid. Note that with either this Irregular Grid or the Regular Grid, polygons can be used to exclude or include specific complex shapes; specify the shape of the polygon from the Krig window (see Define Polygon Outlines later in this chapter).



Block Size
For Block Kriging, this is the size of the block around each point that will be kriged. The local (discretization) grid is placed within this block. For regular grids, block size is defined by the interpolation grid size. For irregular grids, it must be specified here. For Point or Punctual Kriging and for IDW interpolation, block size is irrelevant and is ignored during analysis.

Clear
Clears the worksheet.

Import
The **Import** command brings into the worksheet a file containing interpolate locations. See Chapter 4 (Importing Data from External Files) for instructions on importing.

Other Actions
You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Polygon Outlines (Interpolation Masks)

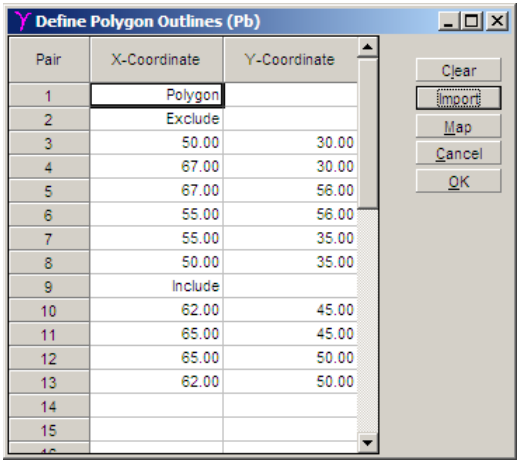
Irregular shapes can be interpolated or excluded from interpolation by defining inclusive or exclusive polygons prior to interpolating. In this window you can define these polygons by entering coordinates of polygon vertices, i.e. x-y points that outline the polygon.

Polygons must be closed, thus they must have at least 3 vertices (3 vertices describe a triangle). You may specify as many vertices per polygon as you like (up to several million), and polygons may be nested within one another as in the example below. However, the vertices for each polygon must define a single, closed polygon, i.e. no segments of the polygon may cross another segment. Also the first and last vertex specified must connect to one another. You can check the shape of the polygon while it is being defined with the **Map** command.

There are two types of polygons, defined by the keyword “Include” and “Exclude”:

- In inclusive polygons, the area within the polygon is interpolated.
- In exclusive polygons, the area within the polygon is not interpolated.

The example below defines two polygons: the first is a 6-sided area that is excluded from interpolation, the second defines an inclusive 4-sided area (rectangle) inside the 6-sided area. You can use the **Map** command to produce a picture of these polygons (see below). The first line in this example (“Polygon”) is unnecessary. Each set of polygon vertices must be preceded by either “Exclude” or “Include”:



Clear
Clear the worksheet.

Import
The **Import** command brings into the worksheet a text file containing vertex locations. The text file can be formatted in a variety of ways with fields separated by either commas or spaces. The default file name extension for polygon interpolation outline files is .int.

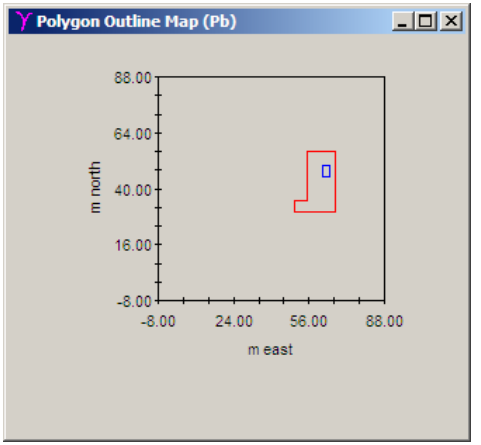
Map
Produces an outline map of each polygon within the larger interpolation grid area. Exclusive polygons appear in red, inclusive in blue. Drawing a map is a useful way to test your polygon coordinates. See the example in the next section.

Other Actions
You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Polygon Outline Map

This window contains an outline map of the polygons used to include and exclude areas from interpolation. It is accessed from the Define Polygon Outlines window, which is part of Interpolation analysis.

The overall area is the interpolation area specified in the Interpolation window. Polygons are defined in the Define Polygon Outlines worksheet. Exclusive polygons appear in red, inclusive in blue:



Graph Actions

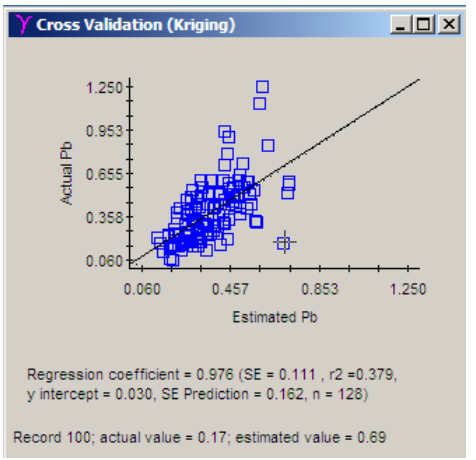
You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS+ window, or via a right-click menu.

Cross-Validation and Jackknife Analysis

Cross-validation analysis is a means for evaluating effective parameters for kriging and IDW interpolations. Two types of validation are provided by GS$^+$: Cross-validation and Jackknife analyses. In cross-validation analysis each measured point in the spatial domain is individually removed from the domain and its value estimated as though it were never there. Then the point is replaced and the next point is removed and estimated, and so on. In this way a graph can be constructed of estimated vs. actual values for each sample location in the domain.

In Jackknife analysis, estimates are compared against measured values for a set of locations different from those used as input data. Before performing jackknife analysis you must specify the jackknife data in a worksheet that appears when you press the Define command.

Cross validation and jackknife analysis are not available for Conditional Simulation.



Each point on the cross-validation and jackknife graph represents a location in the input data set for which an actual and estimated value are available. Information about individual points is provided at the bottom of the screen; points are displayed by placing the cursor on them. In the case above, the cursor was placed on the point representing record 100, as noted at the bottom of the window. By right-clicking on the graph you can list the data for all points (see below).

The regression coefficient described at the bottom of the graph represents a measure of the goodness of fit for the least-squares model describing the linear regression equation. A perfect 1:1 fit would have a regression coefficient (slope) of 1.00 and the best-fit line (the solid line in the graph above) would coincide with the dotted 45-degree line on the graph. The standard error (SE = 0.162, above) refers to the

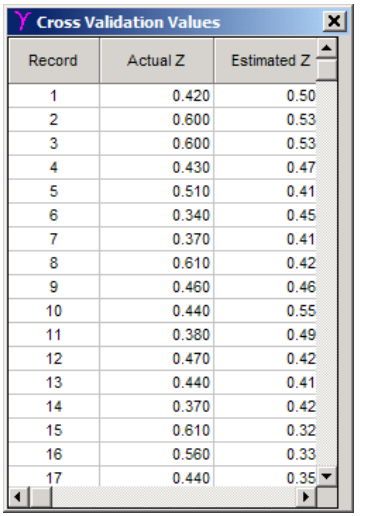
standard error of the regression coefficient; the r2 value is the proportion of variation explained by the best-fit line (in this case 37.9%; it is the square of the correlation coefficient); and the y-intercept of the best-fit line is also provided. The SE Prediction term is defined as $SD \times (1 - r^2)^{0.5}$, where SD = standard deviation of the actual data (the data graphed on the y-axis).

If you have chosen to transform the input data without a backtransformation (see the Data Summary window), then the Estimated Z values will appear very different from Actual Z values regardless of the integrity of the interpolation system.

Cross-Validation and Jackknife Values

In cross-validation analysis each measured point in a spatial domain is individually removed from the domain and its value estimated via kriging as though it were never there. In Jackknife analysis, estimates are compared against measured values for a set of locations different from those used as input data. Results can be graphed (see above) as well as listed in this read-only worksheet table.

Note that Estimated Z values may or may not be backtransformed depending on settings in the Data Summary worksheet. If you have chosen to transform the input data without backtransforming, then the Estimated Z values will appear very different from Actual Z values regardless of the integrity of the interpolation system.



Record	Actual Z	Estimated Z
1	0.420	0.50
2	0.600	0.53
3	0.600	0.53
4	0.430	0.47
5	0.510	0.41
6	0.340	0.45
7	0.370	0.41
8	0.610	0.42
9	0.460	0.46
10	0.440	0.55
11	0.380	0.49
12	0.470	0.42
13	0.440	0.41
14	0.370	0.42
15	0.610	0.32
16	0.560	0.33
17	0.440	0.35

Table

The record number refers to the actual record in the Data Worksheet. For each record successfully kriged the actual Z value is presented in the middle column and the estimated Z value is provided to its right.

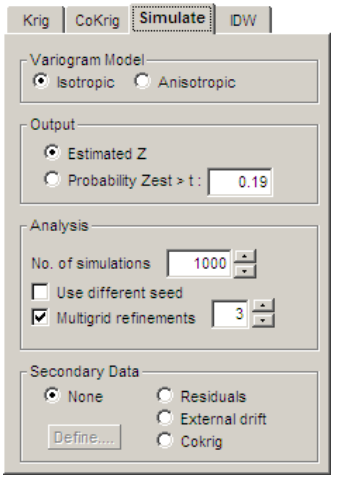
Other Actions

You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or via a right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or use the Data | Change Decimals menu command). Clicking on the top of a column will **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Chapter 10
Conditional Simulation

Conditional simulation is an advanced interpolation technique whereby Z estimates are based on a form of stochastic simulation in which measured data values are honored at their locations. Other interpolation methods, including kriging and IDW, will smooth out local details of spatial variation, especially as interpolated locations become more distant from measured locations. This can be a problem when you are trying to map sharp spatial discontinuities such as contamination hotspots or fault lines. GS$^+$ uses a sequential Gaussian simulation method.

The Interpolation Window contains the Simulate tab:



Variogram Model
Variogram models for isotropic and anisotropic variograms are chosen with the Model command in the Autocorrelation Analysis window. Here in the Simulate window you can specify whether to use the isotropic or anisotropic model for the variogram used in the kriging system.

Output
Two types of output are available with conditional simulation: either a) the estimated Z value and its variance (either estimation variance or standard deviation), or b) the probability that the estimated value for that location is greater than some threshold value *t*.

- An individual Z estimate and its standard deviation is the mean and standard deviation of *n* simulations for a specific location. Thus the number of simulations will affect both of these values.

- Probability is the proportion of simulations for a specific location for which the estimated value is greater than *t*. Thus this value will also be affected by the number of simulations.

Analysis tab

Number of simulations. You may choose any number of simulations >0. Keep in mind that choosing too few simulations will produce a map that is very rough, although a large number of simulations may require substantial cpu time. See later in this chapter for examples. The default number of simulations is specified in the Preferences dialog.

Use different seed. If this box is checked each simulation will use a different random number seed. This will slow the analysis somewhat and make each simulation unique to a specific run. The default is specified in the Preferences dialog. In either case a different random path is used for each simulation.

Multigrid refinements. Checking this box forces the analysis to follow a stepwise procedure when simulating interpolation nodes. In the first step a coarse grid is used to allow the influence of large-scale variogram structure; in subsequent steps the search neighborhoods are smaller. This avoids the need for extensive conditioning. The default is specified in the Preferences dialog.

Secondary data

Collocated secondary data can be used to refine estimations. Collocated means that a secondary data value is available *for every grid node to be estimated* (not simply for every primary data value, as is the case for regular cokriging).

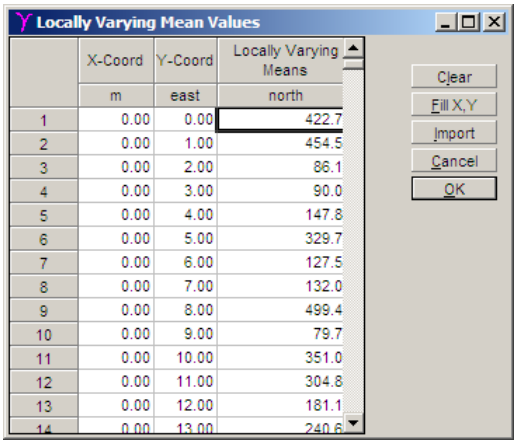
Residuals. This is the same as simple kriging with a locally varying mean. Press the **Define** button to define the residuals to be used.

External drift. Press the **Define** button to define the drift values.

Collocated cokriging. Press the **Define** button to define the covariate values.

Secondary Data for Simulations

Secondary data for conditional simulation are defined in the Simulation Secondary Data window. Secondary data must be available *for every grid node to be estimated*. This window is accessed from the Secondary data tab of the Simulate tab:



	X-Coord	Y-Coord	Locally Varying Means
	m	east	north
1	0.00	0.00	422.7
2	0.00	1.00	454.5
3	0.00	2.00	86.1
4	0.00	3.00	90.0
5	0.00	4.00	147.8
6	0.00	5.00	329.7
7	0.00	6.00	127.5
8	0.00	7.00	132.0
9	0.00	8.00	499.4
10	0.00	9.00	79.7
11	0.00	10.00	351.0
12	0.00	11.00	304.8
13	0.00	12.00	181.1
14	0.00	13.00	240.6

Clear
Clear existing data from worksheet.

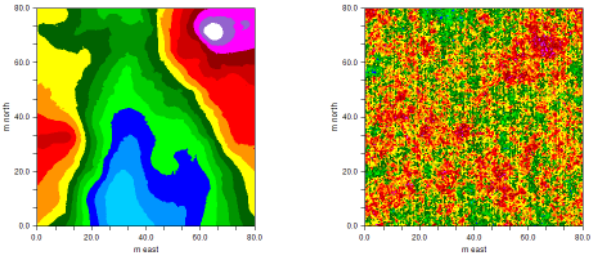
Fill X,Y
Fill the X and Y Coordinate columns with calculated values based on the regular interpolation grid defined in the Interpolation window. If a regular grid is not defined in the interpolation window this button is not available.

Import
Import values from an external text file. Each record of the file must contain a separate x, y, and secondary data value separated by commas, tabs, or spaces. See Chapter 4 (Importing Data from External Files) for instructions on importing.

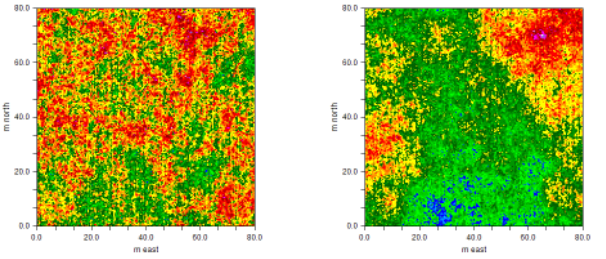
Other Actions
You may **Print, Copy,** or **Export** the contents of the worksheet using the menu commands of the main GS$^+$ window, or the right-click menu. You may also change the **Decimal Places** reported by highlighting a column and pressing the Increase or Decrease Decimals icon (or the Data | Change Decimals menu command). Click on the top of a column to **Sort** the worksheet based on the column selected in alternating ascending or descending order. You may also change **Column Widths** by placing the cursor over the line between two columns and dragging to a new location.

Number of Simulations

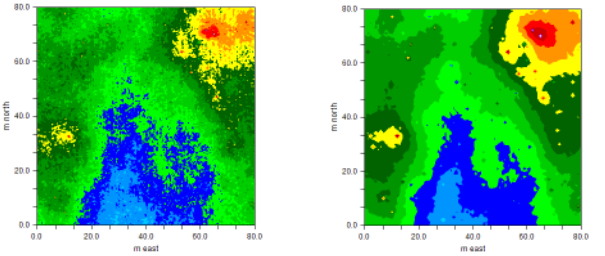
In conditional simulation the number of simulations used to produce estimates of Z can strongly affect the outcome of the interpolation. Below are maps of an 80 x 80 m grid interpolated at a density of 0.5 m, with different numbers of simulations. Below left is a map of block kriged data; on the right is the same data interpolated by conditional simulation (n = 1 simulation).



Additional simulations improve clarity; the map below left represents n = 10 simulations, below right n = 100 simulations:



And still more simulations resolve further detail. The map below left represents n = 1,000 simulations; below right n = 10,000 simulations:

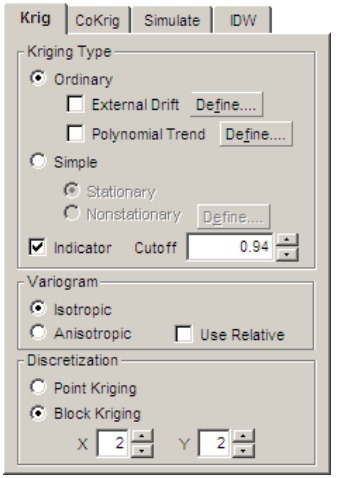


You may wish to set the number of simulations high enough that additional simulations reveal little further resolution.

Chapter 11
Kriging and Cokriging

Kriging provides a means of interpolating values for points not physically sampled using knowledge about the underlying spatial relationships in a data set to do so. Variograms provide this knowledge. Kriging is based on regionalized variable theory and provides an optimal interpolation estimate for a given coordinate location. GS+ performs several types of kriging.

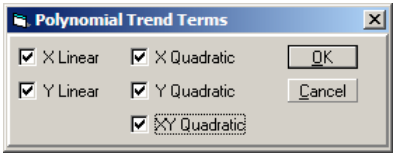
The Interpolation Window contains the Krig tab:



Kriging Type

GS+ provides both ordinary and simple kriging. Ordinary kriging, the most commonly used type of kriging, assumes a constant but unknown mean that may fluctuate among local neighborhoods within a study area. In ordinary kriging the sum of kriging weights equals to one.

Ordinary kriging can be performed with external drift or with a polynomial trend. Kriging with a trend (also called universal kriging) employs a prior trend model, which is defined as a smoothly varying deterministic function. Five different polynomial trend terms can be included in the model: X linear, Y linear, X quadratic, Y quadratic, and XY quadratic; use the Define command to specify which terms to include:



Kriging with an external drift is an extension of kriging with trend. The trend model is limited to two terms $m(u)=a0+a_1f_1(u)$ where m is the mean value for estimation neighborhood u and $f_1(u)$ is set to the secondary (external drift) variable. Use the Define command to assign a drift term to every interpolation grid point.

Simple kriging assumes the expectation of the random field to be known, and relies on a covariance function. Stationary simple kriging has a constant mean, which in GS$^+$ is calculated as the mean of the input data Z. In non-stationary simple kriging, a locally varying mean is specified for every interpolation grid point using the Define command.

For continuous variables, indicator kriging estimates the probability that the interpolation point is greater than a particular cutoff value, specified by the user.

Variogram

Variogram models for isotropic and anisotropic variograms are defined and chosen using the Model command in the Autocorrelation Analysis window. Here in the Kriging window, you can specify whether to use the isotropic or anisotropic model. You can also choose to use a relative variogram, in which the nugget and structural components are rescaled from 0 to 1.0.

Discretization grid

Choose either Point or Block kriging in this section. Your choice should be made on the basis of sampling design and variate characteristics. If samples were taken to represent an area around the actual sample point (e.g. if samples from an area around the sampling coordinate were composited before analysis), then block kriging may be more appropriate than punctual. If samples were taken to represent point values, then punctual kriging will be more appropriate.

For block kriging, you must define the discretization grid (also called the local grid), which is placed around the interpolation point when kriging. The interpolation estimate for that point is based on the mean value of estimates for each of the discretization grid points. You may specify the number of discretization grid points to use in the X and Y boxes; 2 points in the X direction and 2 points in the Y direction means that 4 discretization points will be used.

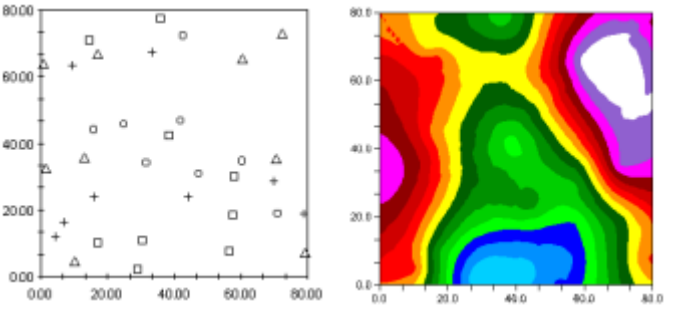
The size of the discretization grid depends on the interpolation grid. For a regular grid, the size is equivalent to the X and Y direction distance intervals, respectively (see Defining a Regular Interpolation Grid in Chapter 9). For an irregular grid, the size is specified in the Define Interpolation Grid worksheet (see Block Size under

Defining an irregular Interpolation Grid in Chapter 9).

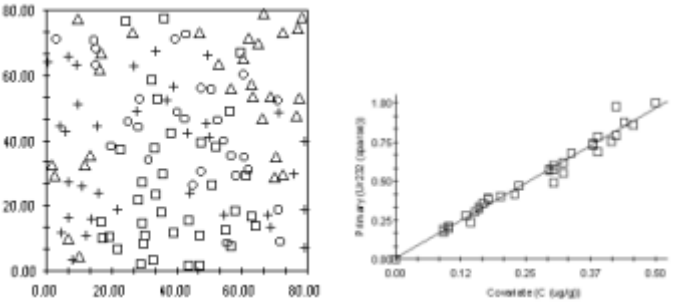
Cokriging

Cokriging is an interpolation technique that allows one to better estimate map values if the distribution of a secondary variate is known. The secondary variate (also called a covariate or Z2) is sampled from the same locations as the primary variate Z, and also from a number of additional locations. If the primary variate is difficult or expensive to measure, then cokriging can greatly improve interpolation estimates without having to more intensely sample the primary variate.

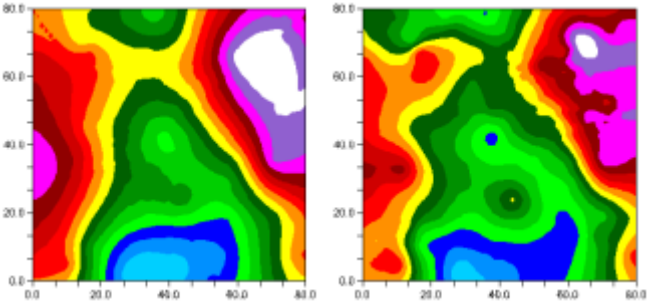
Consider the following hypothetical example. After a radioactive spill, plutonium was sampled across an 80 x 80 m area at a sample density indicated by the quartile plot below. Block kriging (following variography) resulted in the adjacent map:



Soil carbon, easier to measure than Uranium, was sampled at the same locations as Uranium and additionally at another 60 locations as noted in the quartile map below left. Regression of carbon against plutonium showed that the variates were highly correlated (right), suggesting that cokriging might improve the map of plutonium.



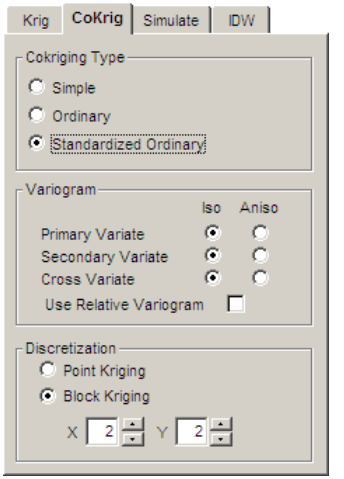
Using carbon as a covariate to produce a cokriged map of plutonium results in the below-right map, plotted next to the original kriged map. Note the substantial improvement in the definition of contour (isoline) differences, especially in the upper right quadrant of the map where the uranium was sampled most sparsely:



The kriging estimate is based not only on distance to nearby sample locations for Z and the variogram for Z, but also distance to nearby sample locations for Z2, the variogram for Z2, and the cross-variogram for Z x Z2. This can provide a more robust estimate of Z at unsampled locations if Z and Z2 are sufficiently correlated.

Prior to cokriging you must a) define a covariate in the Data Worksheet Field Assignment dialog, b) perform semivariance analysis (including variogram modeling) for the primary variate Z, for the covariate Z2, and c) for the cross variate Z x Z2. You will also want to check that the covariate is in fact correlated with the primary variate in the Data Summary – Covariate window.

Once the three variograms are modeled, you can choose the CoKrig tab in the Interpolation window:



Cokriging Type

GS+ provides three types of cokriging: Simple, Ordinary, and Standardized Ordinary.

Simple cokriging places no constraints on the weights applied to the measured values during interpolation.

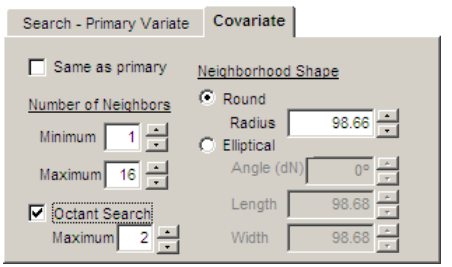
Ordinary cokriging sets the sum of weights applied to the primary variate to one, and the sum of weights applied to the covariate to zero. This limits the influence of the covariate relative to other cokriging types, so may not be preferred.

Standardized ordinary cokriging recalculates the covariate to have the same mean as the primary variate, and constrains all weights to sum to one. This is often the preferred cokriging method and is the default cokriging type in GS$^+$.

Variogram Model Type
Variogram models for isotropic and anisotropic variograms are defined and chosen using the Model command in the Autocorrelation Analyses windows. There is a separate autocorrelation window for the primary variate Z, for the secondary variate Z2, and for the cross variate Z x Z2. Here in the Cokriging window, you can specify whether to use the isotropic or anisotropic model for each of these variograms. You can also choose to use a relative variogram, in which the nugget and structural components are rescaled from 0 to 1.0 for each of the three models.

Discretization grid
Choose either Point or Block cokriging in this section. The choices in this section are the same as for Kriging, above.

Search Neighborhood – Covariate
The search strategy for covariate Z2 values in cokriging can be different from that for the primary variate Z values. In the Search Neighborhood Covariate tab you can specify whether search parameters should be the same as for the primary variate Z or different. See Search Strategy in Chapter 9 for further information about the options available:



Chapter 12
Inverse (IDW) and Normal (NDW) Distance Weighting

Inverse Distance Weighting (IDW) and Normal Distance Weighting (NDW) are interpolation techniques in which interpolated estimates are made based on values at nearby locations weighted only by distance from the interpolation location. Neither IDW nor NDW make assumptions about spatial relationships except the basic assumption that nearby points ought to be more closely related than distant points to the value at the interpolate location. IDW applies stronger weights to nearby points than does NDW. The formula used for Inverse Distance Weighting is:

$$Zest_j = \Sigma [z_i / (h_{ij} + s)^p] / \Sigma [1 / (h_{ij} + s)^p]$$

where

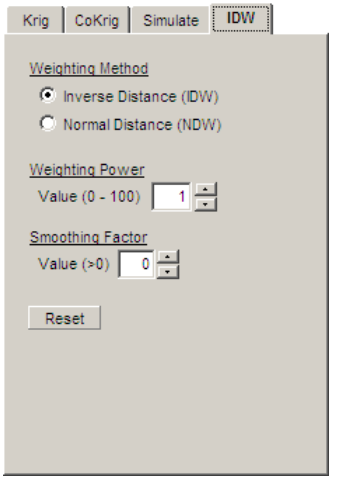
$Zest_j$ = estimated value for location *j*
z_i = measured sample value at point *i,*
h_{ij} = distance between $Zest_j$ and z_i;
s = smoothing factor; and
p = weighting power.

The formula used for Normal Distance Weighting is the same as that for IDW except that the denominator [distance plus smoothing factor] rather than [inverse of distance plus smoothing factor].

The weighting power p defines the rate at which weights fall off with h_{ij}, the distance between the interpolated and sample locations. A value of 1-5 is typical.

The smoothing factor *s* reduces the likelihood that any one sample value will overly influence an estimated value for a given interpolation location. IDW is an exact interpolator, so where an interpolation location *j* coincides with a sample location *i*, $Zest_j = z_i$ and a sharp “peak” or “valley” may result; setting $s > 0$ reduces this peaking effect when it occurs.

The Interpolation Window contains the IDW tab:



Inverse vs. Normal Distance Weighting
Choose either inverse distance weighting or normal distance weighting. Inverse weighting applies stronger weights to nearby points than does Normal Distance Weighting.

Weighting Power
The weighting power defines the rate at which weights fall off with distance between the interpolated and sample locations. A value of 1-5 is typical.

Smoothing Factor
The smoothing factor s reduces the likelihood that any one sample value will overly influence an estimated value for a given interpolation location.

Reset
Restore default values to the IDW method.

Chapter 13
Interpolation Output File Formats

GS+ Format Output Files

The GS+ format for Krig (.krg), IDW, and Conditional Simulation output files (which is also the GS+ format for Map input files) contains a number of header records containing information about the file, followed by XYZ-style records for each interpolated points. Each data record contains fields for an x-coordinate value, a y-coordinate value (except 1-dimensional transects do not contain a value for the y-coordinate), an estimated z value for that x-y location, an estimation standard deviation for the estimated z-value (except IDW files), and the number of pairs used in the interpolation for that x-y location. Missing values are denoted by whatever value is specified in the Preferences window. Note that this format is different from the GS+ Input File format.

- The following listing is for the first 6 coordinates for a standard GS+ krig output file. A GS+ IDW Output file would look identical except that there would be no field for Estimated Standard Deviation in the data records.

```
GS+ Output: Block Kriging Interpolation File v7.0
Set: Field 54, Second Tier
Dimensions: 2; Interval source: Calculated
Interpolation interval (x; y): 1.3333; 1.3333
X-coor: m east
  range: 0.00 - 80.00
Y-coor: m north
  range: 0.00 - 80.00
Z-est:  Pb
  range: 0.151 - 0.813
Z-sd:
  range: 0.0000 - 0.3655
Mean Z-estimate (sd): 0.391 (0.0186)
Valid N: 3721; Missing N: 0; Missing Value Indicator: -99
X-Coordinate    Y-Coordinate    Z-Estimate      EstStdDev       n
0.00    0.00    0.431   0.3608  16
0.00    1.33    0.458   0.3486  16
0.00    2.67    0.462   0.3329  16
0.00    4.00    0.466   0.3178  16
0.00    5.33    0.470   0.3034  16
0.00    6.67    0.474   0.2900  16
```

ArcView® and ArcInfo® Format

The ArcView® Ascii (.asc) format for Krig output files (which is also the Arc/Info® format for Map input files) contains a number of header records containing information about the file, followed by a string of z values for each row of the interpolated grid beginning in a particular corner. Unlike GS+ format files, the standard deviation of the estimate and the number of neighbors used for interpolation are NOT included in this format. Also for this format, the x and y interpolation intervals must be the same (you can set them to be the same from the Interpolation Grid dialog window). This format is compatible with ESRI's ArcInfo Geographic Information System. Data in this format can be imported directly into ESRI's ArcInfo™ or ArcView™. Note that this format is different from the ArcView XYZ Input File format described earlier, which is an ArcView output format that can be read by GS+.

- The following listing is for the first 200 or so coordinates in a standard Arc/Info® or ArcView® format krig output file:

```
ncols         376
nrows          253
xllcorner     -185556.375
yllcorner     -127261.5234375
cellsize      1009.975
NODATA_value  -999
354 314 309 301 286 264 305 306 285 285 268 314 339 306 266 248 252
240 256 282 277 289 269 285 277 258 256 282 268 246 250 249 245 266
262 287 295 335 325 323 359 369 361 357 394 450 407 409 382 409 394
486 510 502 516 546 531 542 572 579 586 594 522 654 550 615 678 709
616 533 430 576 507 638 778 769 456 432 576 635 778 721 769 869 563
589 640 588 770 833 874 1019 965 933 621 795 1131 1044 899 1072
1112 880 910 1069 1088 1071 965 975 902 800 543 350 330 376 310
384 418 352 263 205 202 200 200 201 200 198 201 201 201 201 201 201
694 695
```

To convert this file to an ArcGIS™ grid file, use ArcToolbox™, either the light version that comes with ArvView™ or the full version that comes with ArcInfo™:

a. Open ArcToolbox™

b. Choose the Import to Raster tool

c. Specify the GS+ output file name, the grid type (integer or floating point), the name of the grid output file, and press OK.

d. You may now open the file in ArcMap™.

Surfer® Grid (.GRD) Format

The Surfer® Grid (.grd) Krig Output format (or Map Input format) has a short header area that provides information needed for mapping, and the data is then written as a continuous stream of Z-estimates beginning from a specific corner of the interpolation grid. The standard deviation of the estimate and the number of neighbors used for interpolation are NOT included in this format. This format is compatible with Golden Software's Surfer® mapping program. Note that the Surfer® Grid format is different from the Surfer® XYZ format used for input files (see Chapter 4).

- The following listing is for the first 100 or so coordinates in a standard Surfer® Grid format file. The first record is a required line of code letters that identifies the file as a Surfer® grid file. Record two contains the number of grid lines along the x axis and y axis, respectively. Record 3 contains the minimum and maximum values of the x-coordinate. Record 4 contains the minimum and maximum values of the y coordinate. Record 5 contains the minimum and maximum values for the z-variate. Records 6+ contain the z values organized in row order. Within each row the y-coordinate is constant, and the grid row 1 corresponds to the lowest y-coordinate value and the last row corresponds to the highest y-coordinate. Within each row the z-variate values are ordered from low x-coordinate to highest x-coordinate.

```
DSAA
376   253
-185556.375       194194.225
-127261.523       128262.152
81.000            4469.000
578 746 1097 1251 581 312 229 229 235 278 292 357 490 997 1411
1397 1373 1373 1370 1061 942 1003 1288 1014 785 963 934 797 703
924 1035 1071 1257 1775 1800 1520 1460 1520 1278 1098 915 645 495
450 378 339 342 425 613 1084 1292 1701 1682 1996 2280 2107 1662
1553 1418 1162 1225 1266 1735 1674 1550 1800 2015 2240 2158 2328
2262 2173 2046 1688 1849 1880 2023 2320 2029 2140 2240 2650 2465
2369 2180
```

GSLib Output (.out) Format

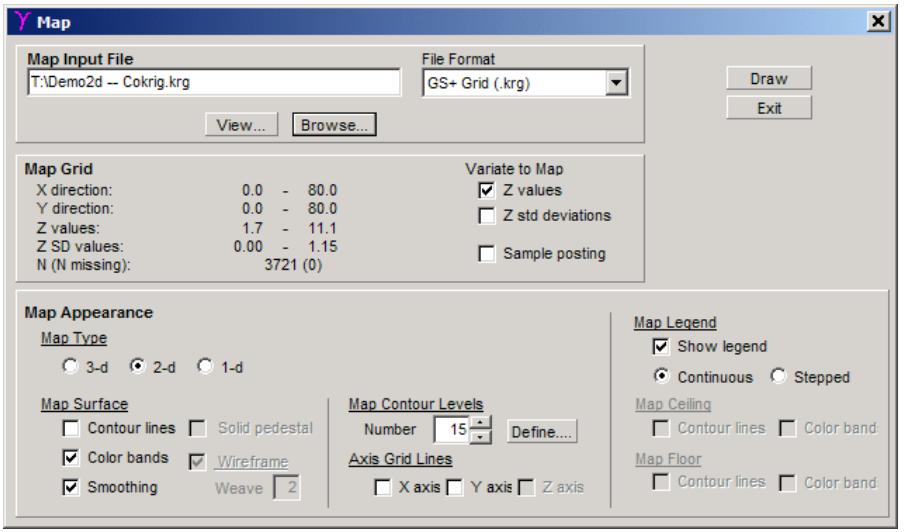
The GSLib (.out) interpolation output format has a short header area that provides information for mapping. The first header record (shown on 2 lines below because of page width limitations) contains x and y coordinate information needed for mapping. The second record states the number of fields per data record, and the next records contain the names of the fields. Following the header records are records that contain a value for the Z-estimate and the estimation variance for that point (or estimation standard deviation if specified in the Interpolation Window). Data records begin in the NW corner of the grid.

- The following listing is for the first few coordinates in a standard GSLib format file. The first record (shown on 2 lines because it is too long to display on one line) identifies the file as a GSLib output file and provides coordinate spacing information. Record two states that there are 3 fields per record, and records 3-5 provide the names of the fields.

```
GS+ Output: Punctual Kriging Interpolation File v9.0; GSLib format: nx=3;
xmin=5.0000; xsiz=10.0000; ny=6; ymin=5.0000; ysiz=10.0000
3
Z-Estimate
EstimationVariance
Neighbors
3.98        6.7825   8
0.34        5.1996   8
1.45        4.5482   8
3.14        5.2250   8
7.67        4.8409   8
0.61        6.2536   8
1.05        4.8234   8
0.86        5.1885   8
3.41        3.9369   8
1.03        5.7149   8
1.44        5.7889   8
2.14        6.2928   8
1.53        6.2740   8
1.15        4.7431   8
0.98        7.2801   8
2.29        9.9600   8
2.28        9.9919   8
2.10        9.9002   8
```

Chapter 14
Mapping

GS+ produces 2-d and 3-d maps of spatial data following interpolation. The data to be mapped come from kriging or IDW analysis, and are thus contained in interpolation output files. Maps can be displayed in a variety of ways with a variety of different contouring schemes, 3-dimensional maps can be rotated on the fly, and both 2-dimensional and 3-dimensional maps can be zoomed to more closely view a transition or other map feature. Additionally, sample postings (original data locations) can be displayed, and estimation standard errors can be mapped for input files that have been saved in the standard GS+ format.



Map Input File

- Select – allows you to choose the file with the data to be mapped. The default file extension will correspond to the selected file format (e.g. .krg, .blk, and .pun for GS+ format), although a file with any extension can be selected.
- View – view the selected file in the View File window.
- Format – specify the format to use to read the input file. Files with a .krg, .idw, .blk, or .pun extension are assumed to be in GS+ format. The format specified in this field will override any assumed format:

- GS+ Krig format (.krg) – in this format a header area defines the interpolation grid, variate names, and other information about the file needed to initiate mapping later, and the data records include for each X and Y Coordinate location that is kriged the interpolation or Z-estimate, the standard deviation of the Z-estimate, and the number of neighbors that were used to make the estimate. See detailed example in Chapter 11.
- Surfer® Grid format (.grd) – in this format a short header area defines information needed for mapping, and the data is written as a continuous stream of Z-estimates beginning from a specific corner of the interpolation grid. The standard deviation of the estimate and the number of neighbors used for interpolation are NOT included in this format. This format is compatible with Golden Software's Surfer mapping program. Note that this format is not the same as the Surfer XYZ Input file format. See detailed example in Chapter 11.
- ArcView® Format (.asc) – this is similar to the Surfer format but the header area is formatted differently and the Z-estimates are written in a pattern that begins from a different corner of the interpolation grid. The standard deviation of the estimate and the number of neighbors used for interpolation are NOT included in this format. Also for this format, the x and y interpolation intervals must be the same (you can set them to be the same from the Interpolation Grid dialog window). This format is compatible with ESRI's Arc-Info Geographic Information System. See detailed example in Chapter 11.
- GSLib Format (.out) (.out) – this format is similar to the GeoEas input format. A long first record contains coordinate interval information that is read by GS+ when mapping GSLib files. The second record of the file specifies three header records named "estimate," "estimation variance" (or "standard deviation"), and "neighbors." Records that follow are in the same order as for Surfer®, but there are three fields per record (estimate, estimation error, and number of neighbors) rather than just one field (the Z-estimate). See a detailed example in Chapter 11.

Map Grid

X direction – the range within the file for values in the x (horizontal) direction.

Y direction – the range within the file for values in the y (vertical) direction; for 1-dimensional data sets this field is blank.

Z values – the range of estimated Z values in the input file.

Z SD values – the range of estimation Z standard deviation values in the input file; only when input files are in a GS+ format that contain both Z values and Z standard deviation values will this box be selectable;

N (N missing) – the number of valid records in the file and (in parentheses) the number of missing values. The default missing value indicator (e.g. -999.) is specified in the Preferences - General dialog; the active missing value indicator for map files is indicated in the Map Contour Intervals window. Missing values are mapped as transparent regions. Exclusive polygons are mapped as missing values, as is any interpolate location for which there is no kriging solution.

Variate to Map

- Z values – maps the estimated Z values in the file
- Z standard deviations or Z estimation variance – maps the estimation Z variance or standard deviations in the file; only files in the GS+ format contain both Z values and Z standard deviation (or Z estimation variance) values.
- Sample Posting – map only the original sample locations. Requires the presence of a posting file, which has the same name as the input file but a .pos extension. This file is created during interpolation; if it is not present you will not be able to map a sample posting. Coordinate Postings (Quantile Plots) of the active data are viewed through the Data Summary – X,Y Coordinates Tab window, described earlier.

Graph Type

- 2-d – displays a flat, 2-dimensional map of the data
- 3-d – displays a 3-dimensional map of the data. The height of the map can be adjusted with the Edit Graph command of the Map window; the perspective can be adjusted with the Rotate command of the Map window.
- 1-d – displays 1-dimensional data (e.g. a geographic transect or a time series) as an x-y graph. This choice is not available for 2-dimensional data.
- Show legend – displays the contour legend next to the map.

Surface

- Contour lines – draws lines between contour intervals
- Color bands – fills the space between contour lines with different colors; colors can be specified from the Define Contour dialog window by pressing Define Contour Levels, below.
- Smoothing – apply slight smoothing to the data to improve visualization.
- Solid pedestal – for a 3-d map fill in the area beneath the surface with a solid color.

- Wireframe – for 3-d maps, drapes an x-y grid over the surface.
- Wireframe Weave – specifies the density of the wireframe grid. A weave of 0 puts a grid line at every data row and column; a weave of 1 skips one row/column, a weave of 2 skips 2, etc.

Contour Levels

- Number – the number of contour levels to put on the map.
- Define – brings up the Map Contour Intervals dialog window (described in greater detail below) that allows you to set break points for individual intervals and colors for contour bands.

Map Legend

Show legend – displays the contour legend next to the map.

Continuous vs. Stepped – display legend as continuous scale with values next to break points (e.g. between color bands) or display legend as stepped format, with values next to boxed colors.

Ceiling

- Contour lines – project contour lines above the plot surface (3-d maps only).
- Color bands – project color bands onto the ceiling of the plot (3-d maps only).

Floor

- Contour lines – project contour lines under the plot surface (3-d maps only).
- Color bands – project color bands onto the floor of the plot (3-d maps only).

Grid Lines

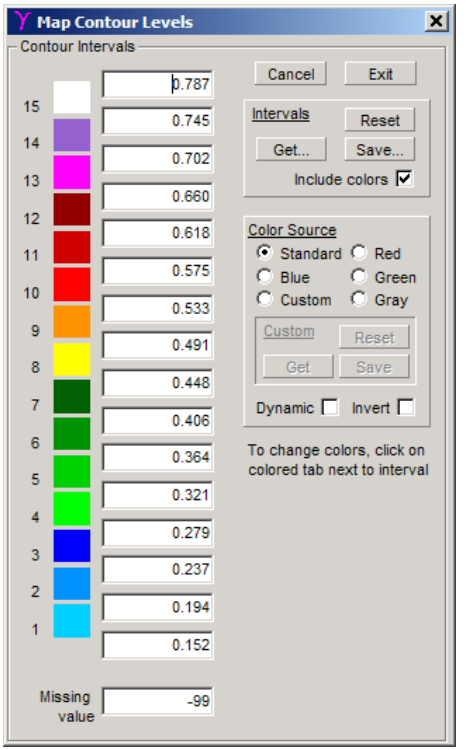
- X axis – place a vertical grid line along the back walls of the 3-d plot x axis.
- Y axis – place a vertical grid line along the back wall of the 3-d plot y axis.
- Z axis – place horizontal grid lines along the back walls of 3-d plots.

Draw

Create the map in a Map Image window.

Map Contour Intervals

This dialog window allows one to specify the break points and colors for the contour intervals used for mapping. The number of contour intervals is specified in the Map window, where access to this Map Contour Levels window is provided.



Contour Intervals

- Color buttons – change the color of a specified interval by clicking the adjacent color button.
- Break points – change the break point between adjacent contour intervals by providing a new value in the space provided. Note that the new break point must be greater than the preceding point and less than the one that follows.
- Missing value – specify the missing value indicator for the data file. This value is usually embedded in the input file as part of the header information, and

is automatically extracted by GS+. This value overrides the missing value indicator specified in the Preferences – General dialog or the map input file.

Intervals

- Get – read contour intervals from a text file; the first value of each record of line of the file is presumed to be a separate interval. The default file extension for Interval files is .lvl.
- Save – save the existing contour intervals to a text file.
- Reset – recalculate intervals based on a regular distribution (even intervals) between the lowest and highest value in the file.

Color Source

- Standard – specifies a default color pattern of red-orange-yellow….. violet.
- Blue – specifies a default color pattern of different shades of blue.
- Custom – specifies a custom color pattern defined below.
- Red – specifies a default color pattern of different shades of red.
- Green – specifies a default color pattern of different shades of green.
- Gray – specifies a default color pattern of different shades of gray.

Color Source Custom

- Get – get a custom color pattern saved earlier. The initial custom color is the same as the standard color source, above.
- Save – save the existing color pattern as the custom color pattern.
- Reset – reset the custom color pattern to the default Standard pattern.

Color Source Dynamic

Expands or contracts the color range when there are fewer than 15 intervals specified. For example, if 10 intervals are specified and the Dynamic box is not checked, the colors for the intervals are the same as they would be if 15 intervals were checked. If, on the other hand, the Dynamic box is checked, the colors for the intervals are adjusted such that interval 10 becomes the color of the original interval 15. Checking this box provides greater color contrast on maps.

Invert

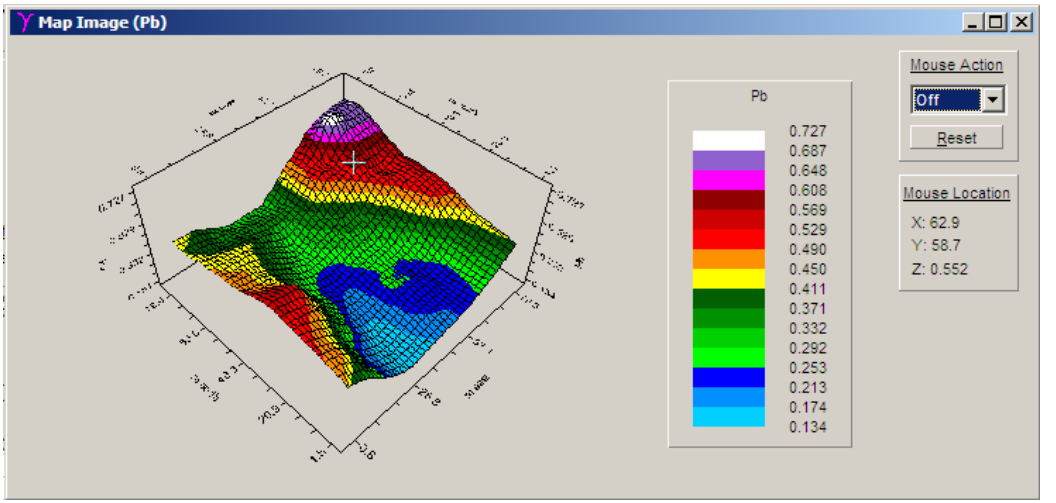
Flips the color range. This allows you to associate the highest interval values with the darkest colors or vice versa.

Cancel

Exit the dialog window without saving changes.

Map Image - 3-d

The 3-d map image is produced by the **Draw** command from the Mapping window.



Mouse Action

- Off – returns the mouse to normal operation.
- Rotate – turns the cursor into a rotator cuff when the left mouse button is pushed, allowing the image to be rotated as desired.
- Move – allows the graph to be moved within the window by clicking the left mouse button and dragging the cursor.
- Scale – shrinks the graph image with the left mouse button.
- Zoom – allows you to zoom in on a graph area by using the left mouse button to define a rectangular zoom area. Within the zoomed area the location of the cursor is noted on the Mouse Location panel.
- Reset – resets the image to the default rotation angle and scale.

Mouse Location

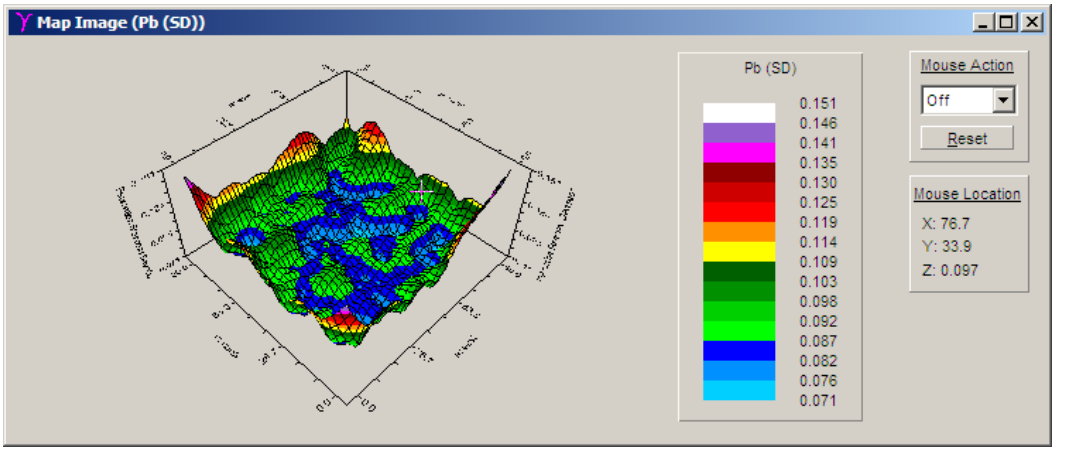
The current cursor location when on the map surface. Units are map units.

Graph Actions

You may **Print, Copy, Edit, Export,** or **List graph values** for either graph using the menu commands of the main GS$^+$ window, or via a right-click menu.

Map Image - 3-d Standard Deviations

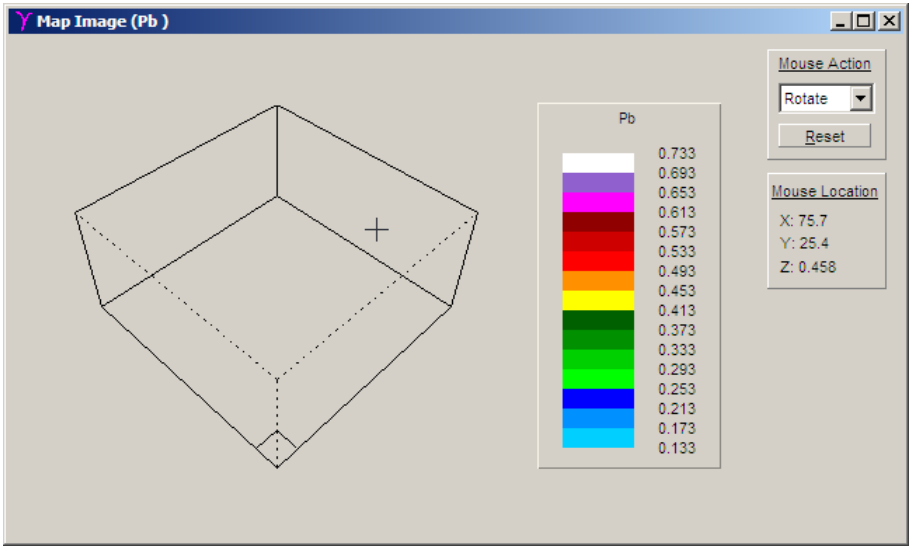
A map of standard deviations can be produced for interpolation files that contain standard deviation or estimation variance values – Krig or Conditional Simulation interpolations that are saved in GS$^+$ format. Use the **Draw** command from the Mapping window.

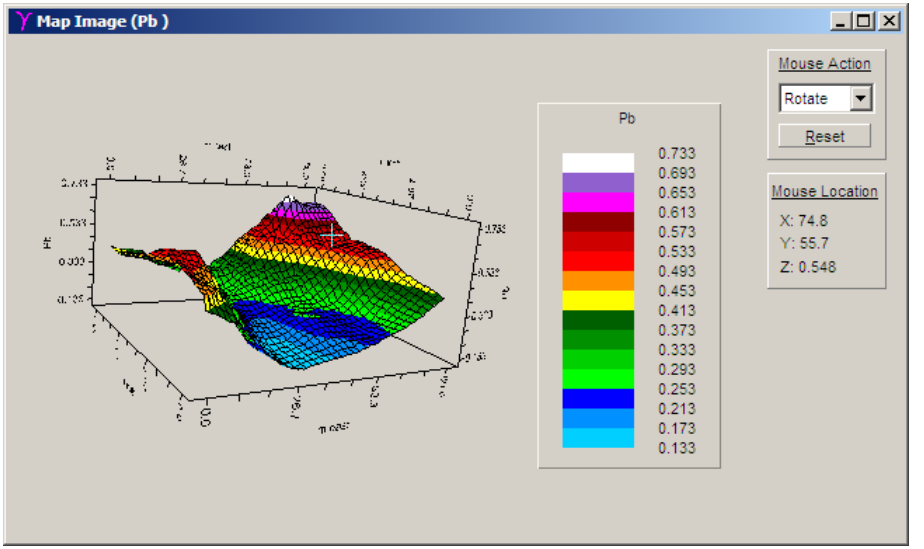


Commands in this Window are the same as for earlier graphs.

Map Image - 3-d Rotation

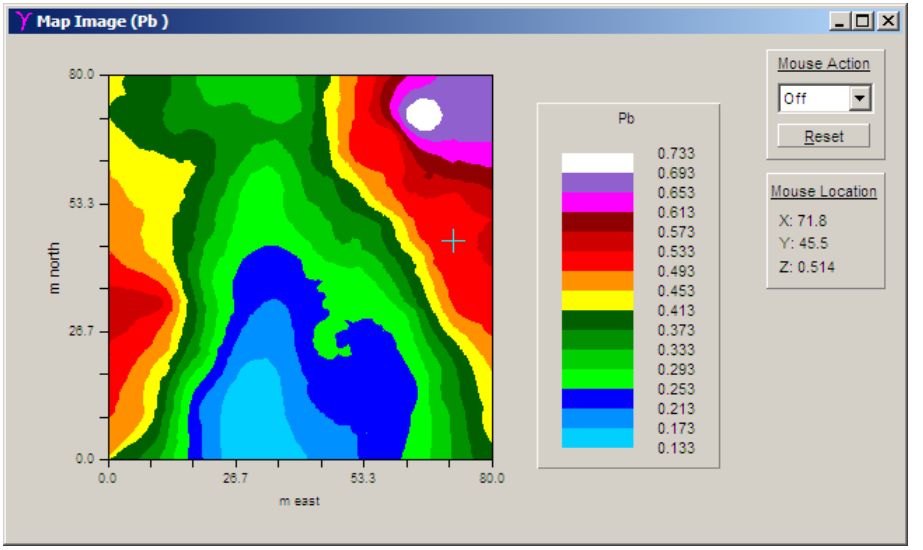
Three-dimensional map images can be rotated to any viewing angle. To rotate, choose Rotate on the Mouse Action panel and hold down the left-mouse button while moving the mouse over the graph. The cursor will be replaced with a 3-d box (as below) of the viewing angle. Release the mouse and the map will be drawn:





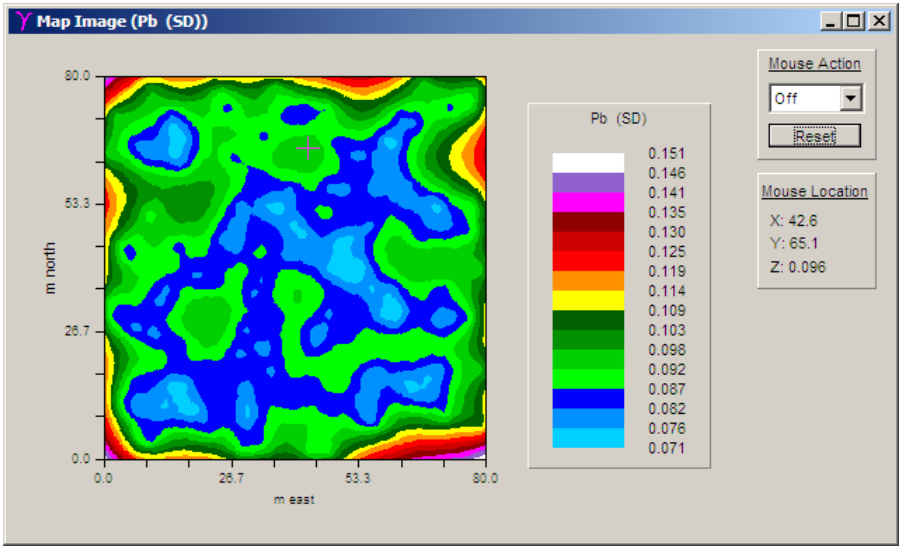
Map Image - 2-d

The 2-d map image is produced when the Graph Type is set to 2-d.



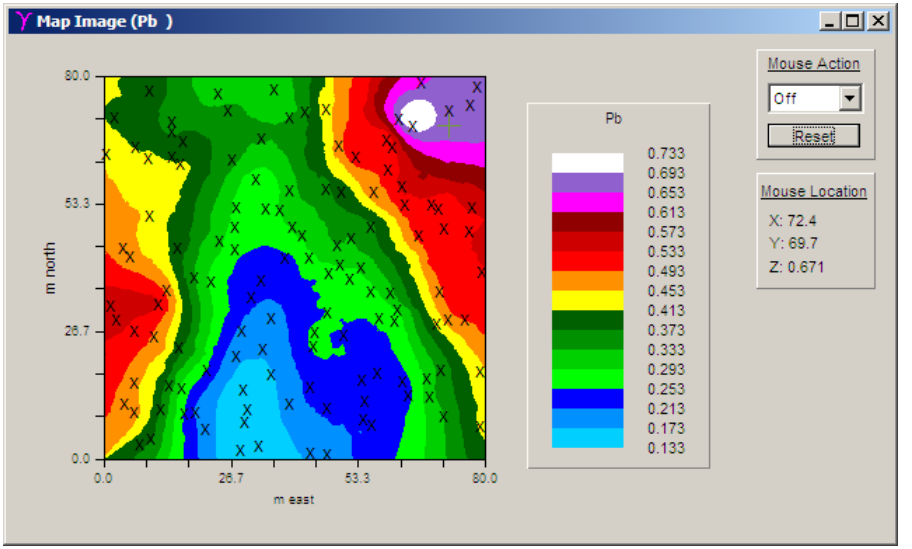
Map Image - 2-d Standard Deviations

A 2-d map of standard deviations can be created from the Map Window.



Map Image - 2-d Sample Posting

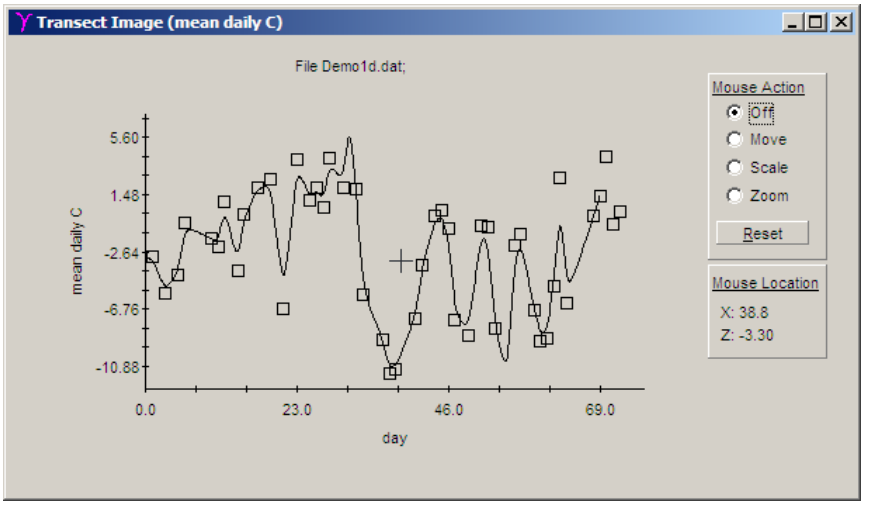
The 2-d sample posting is produced by the **Draw** command from the Mapping window when the Graph Type within that window is set to 2-d and the Posting box is checked. Each "X" symbol in the image below marks an actual sample location as defined by the original X,Y coordinate locations in the Data Worksheet window. Coordinate Postings (Quantile Plots) of the *active* data (which may or may not be the same as the mapped data) are viewed through the Data Summary – X,Y Coordinates Tab window, described earlier.



Commands in this Window are the same as for earlier graphs.

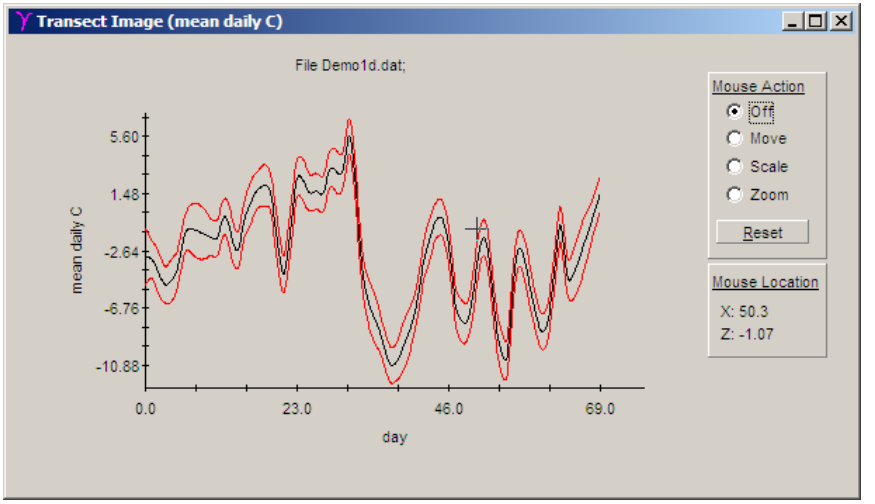
Transects (1-d Maps)

The 1-d transect image with or without sample postings is produced via the Mapping window when the variate to be mapped has only one (X) dimension:



Transect Image with Standard Deviation

A transect with standard deviations can be drawn if interpolation files contain standard deviations – Krig or Conditional Simulation files in GS$^+$ format.



Commands in this Window are the same as for earlier graphs.

Chapter 15
Bibliography

The following references may be useful for those seeking further background about geostatistics and its use in the environmental sciences.

Burrough, P.A. 1981. Fractal dimensions of landscapes and other environmental data. *Nature* 294:240-242.

Burrough, P.A. 1986. Principles of Geographical Information Systems for Land Resources Assessment. Oxford University Press, Oxford.

Burgess, T.M. and R. Webster. 1980a. Optimal interpolation and isarithmic mapping of soil properties. I. The semivariogram and punctual kriging. *Journal of Soil Science* 31:315-331.

Burgess, T.M. and R. Webster. 1980a. Optimal interpolation and isarithmic mapping of soil properties. II. Block kriging. *Journal of Soil Science* 31:333-341.

Cressie, N. 1985. Fitting variogram models by weighted least squares. *Mathematical Geology* 17: 563-586.

Cressie, N. A. C. 1991. *Statistics for Spatial Data*. John Wiley, New York, USA.

David, M. 1977. *Geostatistical Ore Reserve Estimation*. Elsevier, Scientific Publishing Co., Amsterdam, The Netherlands.

Deutsch, C.V. and A.G. Journel 1992. GSLIB Geostatistical Software Library. Oxford University Press, New York.

Goovaerts, P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Griffith, D.A. 1987. *Spatial Autocorrelation: A Primer*. Association of American Geographers, Washington, D.C. 86 p.

Haan, C.T. 1977. *Statistical Methods in Hydrology*. Iowa State University Press, Ames, Iowa.

Isaaks, E.H. and R.M. Srivastava. 1989. *An Introduction to Applied Geostatistics*. Oxford University Press, NY.

Journel, A.G. and C.J. Huijbregts. 1978. *Mining Geostatistics*. Academic Press, New York.

Krige, D.G. 1966. Two dimensional weighted moving average trend surfaces for ore-evaluation. *Journal of the South African Institute of Mining and Metallurgy* 66:13-38.

Krige, D.G. 1981. Lognormal-de Wijsian geostatistics for ore evaluation. *South African Institute of Mining and Metallurgy Monograph Series. Geostatistics I*. South Africa Institute of Mining and Metallurgy, Johannesburg, South Africa.

Mandelbrot, B.B. 1982. *The Fractal Geometry of Nature.* W.H. Freeman, London.

Matheron, G. 1971. The theory of regionalized variables and its applications. *Cahiers du Centre de Morphologie Mathematique, Fontainebleau,* No. 5.

Nielsen, D.R. and O. Wendroth. 2003. *Spatial and Temporal Statistics: Sampling Field Soils and their Vegetation*. Castena Verlag, Reiskirchen, Germany.

Robertson, G.P. 1987. Geostatistics in ecology: interpolating with known variance. *Ecology* 68:744-748.

Robertson, G. P., and K. L. Gross. 1994. Assessing the heterogeneity of belowground resources: Quantifying pattern and scale. Pages 237-253 in M. M. Caldwell and R. W. Pearcy, eds. *Plant Exploitation of Environmental Heterogeneity*. Academic Press, New York, New York , USA.

Rossi, R. E., D. J. Mulla, A. G. Journel, and E. H. Franz. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs* 62: 277-314.

Sokal, R.R. and N.L. Oden. 1978. Spatial autocorrelation in biology. 1. Methodology. 2. Some biological implications and four applications of evolutionary and ecological interest. *Biological Journal of the Linnean Society* 10:199-228.

Trangmar, B.B., R.S. Yost and G. Uehara. 1985. Applications of geostatistics to spatial studies of soil properties. Pages 45-94 *in* N.C. Brady, editor. *Advances in Agronomy Volume 38.* Academic Press, New York.

Vieira, S.R., J.L. Hatfield, D.R. Nielsen, and J.W. Biggar. 1983. Geostatistical theory and application to variability of some agronomical properties. *Hilgardia* 51:1-75.

Webster, R. 1985. Quantitative spatial analysis of soil in the field. Pages 1-70 *in* B.A. Stewart, editor. *Advances in Soil Science Volume 3.* Springer-Verlag, New York.

Webster, R. and M.A. Oliver. 2007. *Geostatistics for Environmental Scientists (Trends in Organizational Behavior).* John Wiley and Sons, NY. 304 pp.

Chapter 16
Glossary

A_0
Range parameter in variogram models. The relationship of A_0 to range A (effective range) depends on the model. See specific model definitions for more information.

A_1
Range parameter for the major axis in anisotropic variogram models. The relationship of A_1 to major range depends on the model. See specific model definitions for more information.

A_2
Range parameter for the minor axis in anisotropic variogram models. The relationship of A_2 to major range depends on the model. See specific model definitions for more information.

Binary Data Record Format
Binary data from many spreadsheet and database programs can be imported directly; if the file can be viewed using the View command then it can be imported. Eligible files include Excel and Lotus 1-2-3 worksheets, dBase, Paradox, Access, etc.

A Character as Missing Value Indicator
Specify the character to be used as a missing value indicator in the field that will appear when a character value is specified in the list box. The example below presumes that % has been specified the missing value indicator, and thus the third value of this record will be read into the worksheet as a missing value:

13.2, 34.5, % , 0.15

Autocorrelation
Autocorrelation is the degree to which a *property is related to itself in time or space.* For example, in a spatial domain values for samples taken close to one another are more likely to be similar than are samples taken farther apart – autocorrelation is a formal measure of this self-similarity.

Backtransformation
When data values are transformed in order to make their distributions more normally distributed, after analysis of the transformed data the output data are customarily (but not necessarily) back-transformed to the original data domain for final reporting. For example, the backtransformation for a ln(z) transform is exp(z-transformed); for a z^2 transform the backtransformation is the square root of the transformed value.

Offset values (if applied) are subtracted from the backtransformed values.

Binary Data Record Format
Binary data from many spreadsheet and database programs can be imported directly. Eligible files include Excel and Lotus 1-2-3 worksheets, dBase, Paradox, Access, and FoxPro.

Block Kriging
Kriging provides a means of interpolating values for points not physically sampled using knowledge about the underlying spatial relationships in a data set to do so. Block kriging provides an estimate for a discrete area around an interpolation point. Punctual or point kriging provides an estimate for a precise point. In environmental work block kriging is usually more appropriate. The block is defined as the rectangular area around a point that is not included in an adjacent block.

Brackets as Column Title Separators
Column or variate names are separated by brackets, e.g.:

<m east> <m north> <Pb (ug/g)> <pH>

C or Structural Variance
C of the variogram model represents spatially-structured variance. Compare to C_0 or nugget variance, which is the portion of the variance not spatially structured.

C_0 or Nugget Variance
C_0 or Nugget Variance is the y-intercept of the variogram model. Nugget variance represents variation not spatially dependent over the range examined.

Characters as Column Title Separators
Column or variate names are separated by a character specified in the field that appears when this option is selected. For example if "#" were specified as the separated:

m east # m north # Pb (ug/g) # pH

Character Separated Values
Values within data records are separated by a specific character, defined in the field below the list box; e.g. if the character “#” is the delimiter:

13.2 # 34.5 # 35.6 # 0.15

Cokriging
An interpolation technique that allows one to use a more intensely sampled covariate in the estimation of values for a related variate. If the primary variate is difficult

or expensive to measure and it is correlated with a more available covariate, cokriging can greatly improve interpolation estimates. See Cokriging Analysis for further information.

Commas as Column Title Separators
Column or variate names are separated by commas e.g.:

m east, m north, Pb (ug/g), pH

Comma Separated Values
Values within data records separated by commas:

13.2, 34.5, 35.6, 0.15

Conditional Simulation
An advanced interpolation technique whereby Z estimates are based on a form of stochastic simulation in which measured data values are honored at their locations.

Covariate
The Covariate is a second Z-variate that covaries with the primary Z-variate. It is used in cokriging.

Data Builds
Data in the GS$^+$ Worksheet is read into the actual arrays that get analyzed during Data Builds. Data builds occur whenever a new data set is read into GS$^+$ (via the Import command) or whenever a field assignment or data value is changed and the Rebuild Button (Data Worksheet Window) is pressed. During data builds missing values are ignored and filters and transformations are applied. Most Data Builds are automatic; you'll need to force a Data Build whenever the Rebuild Button is red (the Data Summary Window will also be empty).

Data Fields
A data field is an individual item on the data record. In the GS$^+$ Worksheet, records appear as rows and fields appear as columns. Prior to analysis you must identify an X-coordinate field, a Y-coordinate field (for 2-dimensional analyses), and the field (column) that corresponds to the Z-variate. For cokriging you will also need to assign a field to hold the covariate (Z2) values. Field assignments are performed by clicking on the top row of the Worksheet to bring up a Field Assignment dialog window.

Data Records
Data Records are records in text files that contain data values. Data records follow header records that contain descriptive information about the file. Data records are imported into the GS$^+$ Worksheet; header records are imported into the Data De-

scription portion of the Worksheet Window.

Decimal Point Missing Value Indicator
In the record below the third value appears as a decimal point and will be read into the worksheet as a missing value:

13.2, 34.5, . , 0.15

Discretization Grid
The discretization grid describes the size of the grid placed around the interpolation point when block kriging. The estimate for the interpolated block is based on the mean value of estimates for each of the discretization grid points. A single discretization point describes punctual kriging. A 2x2 grid averages 4 grid points. Larger discretization grids take longer to interpolate.

Effective Range
See Range.

Geo-referenced Data
Geo-referenced data are data that have been collected from a specific location, i.e. any data for which there is a spatial x,y coordinate associated with each data value.

Header Records
Records at the top of a data file that precede the data records and that contain information about the data records.

Interpolate
An interpolate is an estimated value for a location in a domain based on values of nearby measured locations. In GS^+ interpolation is performed by kriging or by inverse distance weighting methods. Maps are based on interpolations at grid points across the domain.

Kriging
An interpolation method based on regionalized variable theory that provides an optimal interpolation estimate for a given coordinate location.

Lag Class Distance Interval
Each point on a variogram corresponds to the average semivariance for all values in the spatial domain that are separated by a given distance interval *h*. This interval (also called a lag distance class) is a range whose width is specified by the user in the Semivariance Analysis window. A corresponding definition is also relevant to Moran's I, Fractal, Cross-Semivariance, and Madogram Analyses.

Missing Value
A value not present in a record. A data record can contain fields for coordinates (e.g. X, Y) and for multiple Z variates. If any one of the fields used for an analysis is missing, the record is ignored during analyses, i.e. when the active data set is built. In the GS+ Worksheet a value is considered missing if the cell is blank (a permanent missing value) or if the cell contents appear as red italicized text (a temporarily missing value). Cells can be assigned temporarily missing by right-clicking on them; a second right-click returns them to non-missing. In Input and Output files missing values are indicated by characters or values (e.g. -99) defined in the Preferences window.

Model Parameters
See Variogram Model Parameters.

No Header Records
The first record of the file contains numeric records that should be treated as data.

Number of Header Records are Fixed
An input field beneath the list box appears when this choice is made, and this field specifies the number of records at the top of the file to consider descriptive text when reading the file; the first data record is the next record.

Number of Header Records Varies
The number of header records will be assumed to be all records prior to the first all-numeric record. If all records contain an alphanumeric Sample ID field then all records will be assumed to be header records and no data records will be read. In the example below there are 2 header records:

```
File GMD Test Site
<m east> <m north> <Pb> <pH>
34.5   45.6   0.231  5.8
36.7   46.5   0.241  5.9.
```

Numeric Value as a Missing Value Indicator
The indicator value that will appear when a numeric value is specified in the list box. The example below presumes that -99 has been specified the missing value indicator and thus the third value of this record will be read into the worksheet as a missing value:

13.2, 34.5, -99.0 , 0.15

Parameter File
Files that store analysis parameters and data for a particular data set. GS+ parameter files typically have a .par extension.

Polygons
Polygons are Irregular shapes that can be interpolated or excluded from interpolation during kriging. Polygons are defined by at least 3 vertices (coordinate pairs) that define the polygon outline. Polygons can be as complex as continent outlines or as simple as building outlines.

Posting
A Coordinate Posting is a map of coordinate locations. The location of each data point in the active data set can be marked by a symbol corresponding to its relative value.

Punctual Kriging
A type of kriging that provides an estimate for a precise point. Block kriging provides an estimate for a discrete area around a specific interpolation point.

Quotes as Column Title Separators
Column or variate names are separated by quotes. E.G.:

"m east" "m north" "Pb (ug/g)" "pH"

Range
The range (also called effective range) is the separation distance over which sample locations are autocorrelated, i.e. over which there is spatial dependence among sample locations. In variogram models the range is calculated from the range parameter A_0 (isotropic models) or A_1 and A_2 (anisotropic models). See model definitions for further information.

Same as for Data Records
The delimiter for column or variate titles is the same as specified for data records. E.G. if the data records in a file with four data columns are formatted with "#" as a delimiter:

m east # m north # Pb (ug/g) # pH

Sample ID
The Sample ID is text or a numeric value that identifies a particular data record. It is optional. To specify a Sample ID column click on the top row of the data worksheet.

Sill or C_0+C
The Sill of the variogram model (C_0 + C) represents spatially-independent variance. Data locations separated by a distance beyond which semivariance does not change — i.e. after the model asymptote or sill — are spatially independent of one another. Theoretically the sill is equivalent to sample variance.

Space Separated (free format) Values
Values within data records are separated by spaces, e.g.:

13.2 34.5 35.6 0.15

Spaces as Column Title Separators
Column or variate names are separated by spaces; note that this option limits column titles to single words. e.g.:

east north Pb pH

Tab Separated Values
Values within data records are separated by tab characters (denoted by "[tab]" in the example below; this format is often called tab-delimited output by spreadsheet programs, e.g.:

13.2 [tab] 34.5 [tab] 35.6 [tab] 0.15

Tabs as Column Title Separators
Column or variate names are separated by tab characters, e.g.:

m east [tab] m north [tab] Pb (ug/g) <tab> pH

Transformation
A transformation is a formula applied to all Z or Z2 values in a data set, usually to better normalize its distribution. The success of a transformation can be judged by observing its frequency distribution before and after transformation. Values may be backtransformed prior to reporting results.

Regular Lag Interval Classes
The Lag Class Distance Interval defines how pairs of points will be grouped into lag classes. Each point in a variogram or autocorrelogram represents the average semivariance or Moran's I for a single lag class, which is a group of pairs separated by a certain Lag Class Distance Interval, sometimes called a step size. This interval can either be calculated by GS^+, in which case it will be regularly distributed across the active lag distance, or it can be manually set by the user

Variogram Model Parameters
Model parameters for isotropic variograms include terms for nugget variance (C_0), the combined nugget plus structural variance ($C_0 + C$, also called the model sill), and the range parameter A_0 or in the case of anisotropic variograms, A (a function of A_1 and A_2). The range parameter may be different from the effective range, which should be used to compare ranges among models.

What is a valid record
A valid record is any record that contains non-missing values for the coordinate locations (both an X-Coordinate value and a Y-Coordinate value for 2-dimensional domains) AND a non-missing value for the Z-Variate. Missing values appear as blank cells OR as red colored text; use the right mouse button to turn cells with values into cells with missing values and vice versa.

X,Y-Coordinates
X,Y coordinates describe a physical location at which a Z variate is measured. The coordinates are presumed to be in Cartesian space, i.e. with a 0,0 origin that increases for x in an easterly direction and decrease in a westerly direction, and for y increase in a northerly direction and decrease in a southerly (note that values can be less than or greater than 0). Thus a value of 10,20 means 10 units east of the origin and 20 units north.

In GS$^+$, if a Y-Coordinate Column is not specified the data are assumed to be 1-dimensional, i.e. collected along a transect or through time. To specify (or remove) a Y-Coordinate column, click on the top row of the worksheet.

Note also that latitude and longitude are not part of a Cartesian coordinate system because coordinate distances (measured in degrees) represent different physical distances (measured in meters) in different places on the globe. For a rough conversion from latitude and longitude to meters, consider that one second of latitude equals 30.92 meters on the ground; for longitude, calculate the cosine of the latitude then multiply by 30.92.

Z-Variate Data (Z)
The Z-Variate is the variate being analyzed and mapped, e.g. elevation for a topographic map, pH for a map of soil acidity, chlorophyll content for a map of lake productivity, population density for a map of rural population growth, etc. A Worksheet can contain many variates but only one can be analyzed at a time, thus only one column can contain the Z-Variate data. A Z-Variate Column must be defined prior to spatial analysis. To specify a column, click on the top row of the worksheet.

Z2 Variate Data (Z2 or Covariate)
The Z2 Variate or Covariate is a second Z-variate that covaries with the primary Z-variate. It is used in cokriging.

Index

1-d transects, 158
2-d maps, 33, 147, 156
3-d maps, 33, 147, 153
  ceiling, 34
  floor, 34
  grid lines, 34
  map proportion, 33
  Standard deviations, 154
A_0, 79, 80, 82, 83, 84, 85, 167
A_1, A_2, 86, 87, 89, 90, 91, 168
Activation, 7
  activate over the web, 10
  activation code, 7, 8
  activation window, 8
  advanced options, 9
  deactivation, 7, 11
  evaluation mode, 8
  installation ID, 10
  license number, 8
  proxy server, 9
  serial number, 7
  Serial number, 8
  unlocking code, 10
Activation server, 22
Active lag distance, 27, 71
Adjacency matrix, 106
Anisotropy
  axis orientation, 73, 92, 94
  semivariance surface, 73, 74, 94
  Semivariance surface, 92
  surface map, 27
  variogram map, 92, 94
  variogram models, 86
Appending data, 53
ArcView®
  input file format, 46
  map input files, 148
  output file format, 120
  output files, 144
Assign field, 40
Autocorrelation
  active lag default, 27
  autocorrelation window, 71
  correlograms, 100
  covariance analysis, 101
  defaults, 27
  definition, 162
  drift, 99
  fractal analysis, 107
  general relative variograms, 102
  madograms, 97
  menu, 19
  moran's I analysis, 106
  number of lag intervals, 27
  offset, 27
  pairwise relative variograms, 104
  rodograms, 98
  semivariance, 70
  standardized variograms, 96
  subsample data, 27
Axis
  formats, 32
  orientation, 73
Azimuth, 93
Backtransformations, 55, 162
Base input file, 37
Bibliography, 159
Binary data record format, 163
Block
  cokriging, 140
  kriging, 124, 137, 163
  size, 124
C_0, 79, 163
Cartesian (x,y) coordinate system, 15, 169
Ceiling, 34, 150
Character separated values (csv), 163
Clear, 38
Cokriging, 41, 71, 121, 138
Cokriging type, 139

Column, 38
  field assignments, 37, 39
  titles, 47, 163, 164, 167, 168
  widths, 39
Commas as column title separators, 52
Conditional simulation, 117, 121, 132
  colocated cokriging, 133
  external drift, 133
  interpolation defaults, 28
  multigrid refinements, 28
  number of simulations, 28, 135
  residuals, 133
  secondray data, 133, 134
  seed value, 28
Context menu, 23
Contour levels, 34, 150, 151
Convert GS+ DOS File, 35
Coordinate postings, 62
  1-d scattergrams, 65
  definition, 167
  map data, 149, 157
  posting intervals, 66
Coordinates
  assigning, 39
  cartesian, 169
  latitude, longitude, 169
Copying data and graphs, 35
Correlograms, 100
Covariance analysis, 101
Covariate (Z2)
  assignment, 40
  definition, 164, 169
  icons, 16
  regression, 67
  values warning, 41
Cross-semivariance analysis, 20, 95, 139
Cross-validation
  analysis, 121, 129
  values, 131
Cross-variate regression, 67, 68, 69
Cross-variograms, 20, 95, 139
csv files, 18, 35, 52
Cumulative frequency distribution
  graph, 55, 58
  values, 58
Data
  append dialog, 53
  bounding, 42
  builds, 164
  copying, 35
  exporting, 35
  fields, 39, 164
  file import, 44, 52
  filter, 38, 42
  input file format, 45, 47
  menu, 18
  records, 38, 47, 48, 49, 50, 52, 164
  summary window, 62, 67
    regression analysis tab, 67
    x,y coordinates tab, 62
  summary window Z tab, 54
  title, 38
  worksheet, 37
Data file import, 25
Database input files, 44, 51
Deactivate, 22
Deactivation, 7, 11
  activation code, 11
  license number, 11
  proxy server, 11
Decimal places, 18, 24, 26, 39
Define posting intervals, 66
Delete, 18
Descriptive statistics, 55
Discretization grid, 124, 137, 140, 165
Distance interval, 123
DOS GS+ files, 35
Drift, 39, 99
Duplicate values, 43, 55
Edit menu, 17
Effective range, 79, 167, 168
Email, 22
Estimation variance, 120
Estimation variance map, 154
Excel files, 51
Exclusive polygons, 119, 126, 128

Exponential anisotropic model, 89
Exponential isotropic model, 83
Export
  data, 18, 35
  graphs, 17, 35
External drift, 39
  assignment, 40
Field (column)
  assignments, 37, 38, 39
  default assignments, 25
  delimiter, 52
  names, 38
File
  extensions, 35, 75, 120, 143, 144, 145, 146, 148, 166
  import command, 37
  import dialog, 44
  import properties, 51, 52
  view window, 53
File menu, 16
File view window, 53
Filtering data, 38, 42
Floor, 34, 150
Formulas in worksheets, 39
Fractal analysis, 107
Frequency distribution
  graph, 55, 56
  values, 57
Gaussian anisotropic model, 90
Gaussian isotropic model, 85
General relative variograms, 102
GeoEas input file format, 45, 48
Geometric anisotropy, 73
Geo-referenced data, 165
Graph settings
  axis scaling, 31
  axis titles & labels, 32
  colors, 29
  contour details, 33
  footnote, 30
  general, 29
  legend, 30
  title, 30
Graphs
  copying, 35
  data, 19
  exporting, 35
  printing, 35
Grid lines, 34, 150
GS+ Input file format, 45
GS+ Output files, 143
GSLib
  output files, 120, 148
Header records, 47, 48, 49, 50, 52, 165, 166
Help menu, 21
Histogram bars, 32, 56
h-Scattergrams, 108, 113, 116
html files, 18, 35
IDW, 141
  interpolation defaults, 28
  output files, 143, 144, 145, 146
  smoothing factor, 28, 141, 142
  weighting power, 28, 141, 142
Importing
  data, 18
  default file type, 25
  default filename, 25
  file import dialog, 44
  import file command, 37
  spreadsheets, 51
  text files, 45
Inclusive polygons, 119, 126, 128
Indicator kriging, 137
Input file format, 45
  ArcView® XYZ, 46
  ArcView® XYZ, 49
  Custom, 46
  Excel, 51
  GeoEas, 45, 48
  GS+, 45, 47
  spreadsheet and database formats, 51
  Surfer® XYZ, 46, 50
Insert, 18
Insert rows or columns, 39
Installation, 6

Interpolate, 165
Interpolation, 117
  block cokriging, 140
  block kriging, 137
  cokriging, 117, 138
  conditional simulation, 117, 132
  constrain values, 119
  cross-validation analysis, 121, 129
  discretization grid, 137, 140
  distance interval, 123
  draw map after interpolating, 28
  grid, 119
  indicator kriging, 137
  interpolate locations, 125
  interpolate menu, 20
  inverse distance weighting, 117, 141
  irregular grid, 119, 124
  jackknife analysis, 121, 129
  kriging, 117, 136
  kriging with external drift, 137
  kriging with trend, 136
  masks, 119, *See Polygons*
  menu, 20
  nearest neighbors, 120
  neighborhood radius, 121
  normal distance weighting, 117, 141
  number of points, 123
  octant search, 121
  optimize range, 123
  order of output, 28
  ordinary cokriging, 140
  ordinary kriging, 136
  output defaults, 28
  output file formats, 119
  output files, 119
  output standard deviation, 120
  output variance, 120
  polygons, 119, 126
  punctual cokriging, 140
  punctual kriging, 137
  range, 118, 123
  regular grid, 118, 122
  relative variogram, 137, 140
  search neighborhood, 120
  simple cokriging, 140
  simple kriging, 137
  specific shapes, 119
  standard deviation, 28
  standardized ordinary cokriging, 140
  type, 121
  universal kriging, 136
  validation, 121, 129
  variance, 28, 120
  variogram model, 132, 137, 140
  window, 118
  Z estimate boundaries, 122
Interval class, 70
Inverse distance weighting, 117, 121, 141
Irregular interpolation grid, 119, 124
Isotropic variogram models, 79
Jackknife
  analysis, 121, 129
  values, 131
  worksheet, 121
Kriging, 117, 121, 136
  ouput standard deviation, 120
  ouput variance, 120
  output files, 143, 144, 145, 146
  type, 136
  with external drift, 137
  with trend, 136
Lag class distance interval, 27, 70, 71, 72, 75, 165
Latitude, 15, 169
Legend title, 66
License number, 8
Linear anisotropic model, 90
Linear isotropic model, 84
Local grid, 137
Longitude, 15, 169
Madograms, 97
Main menu, 16
Major axis, 86, 93
Map
  2-d image, 156
  2-d sample postings, 157

3-d image, 153
3-d standard deviation image, 154
ArcView® input files, 144, 148
ceiling, 34, 150
contour levels, 34, 150, 151
estimation variance, 149
floor, 34, 150
grid (range), 148
grid lines, 34, 150
GS+ input file format, 143
GS+ input files, 148
GSLib input files, 146
input file formats, 143, 147
legend, 34, 150
mapping in GS+, 147
menu, 20
mouse action, 153
rotate graph, 153
rotation, 155
scale, 153
standard deviation, 149
surface, 149
surface properties, 33
Surfer® input files, 145, 148
zoom, 153
Mapping
draw map after interpolating, 28
interpolation window, 118
Masks, 119
Menus
autocorrelation, 19
context, 23
data, 18
edit, 17
file, 16
help, 21
interpolate, 20
krig, 20
main, 16
map, 20
right-click, 23
window, 21
Minor axis, 86
Missing values
count, 149
definition, 166
missing value indicator, 24, 41, 45, 47, 49, 50, 51, 52, 143, 149, 151
permanent missing values, 39, 41
preferences, 23
temporary missing values, 39, 41, 43
Model parameters, 80, 87
Model results, 77
Moran's I analysis, 106
Mouse action, 64, 93, 153
Multigrid refinements, 28
NDW, 28, 141
Nearest neighbors, 120
Neighborhood radius, 121
Normal distance weighting (NDW), 28, 141
Normal probability distribution
graph, 55, 59
values, 60
Nugget variance, 79, 80, 87, 163
Octant Search, 121
Offset tolerance, 73
Ordinary cokriging, 140
Ordinary kriging, 136
Ouput file formats, 119
Outliers, 108
Outline map, 128
Output files
ArcView® format, 144
ArcView® format, 120
GS+ format, 120, 143
GSLib format, 120, 146, 148
interpolation, 119
Surfer® format, 145
Surfer® format, 120
Pairwise relative variograms, 104
Parameter files, 16, 166
Polygons
definition, 167
exclusive, 119, 126, 128
how to create, 119

inclusive, 119, 126, 128
map, 128
polygon definition window, 126
vertices, 126
Posting intervals, 66
Postings. *See Coordinate postings*
Preferences
analyses, 26
anisotropy surface, 27
autocorrelation Active lag default, 27
autocorrelation defaults, 27
autocorrelation lag intervals, 27
data file import, 25
decimal places, 24
general, 23
IDW smoothing factor, 28
IDW vs. NDW, 28
IDW weighting power, 28
import field assignments, 25
import file type, 25
import filename, 25
interpolation output, 28
missing value indicator, 24
number of subsamples, 28
simulation multigrid refinements, 28
simulation number, 28
simulation seed value, 28
subsample data, 27
variogram offset, 27
Primary variate, 138
Principal anisotropic axis, 93
Printing
graphs, 35
setup, 17
worksheets, 35
Punctual cokriging, 140
Punctual kriging, 137, 167
Quotes as column title separators, 52
Range, 79, 80, 86, 87, 167, 168
Rebuild command, 38
Records. *See Data records*
Regression
cross-validation, 129
regression analysis, 67, 68
regression coefficient, 81, 88
regression values, 69
Regular interpolation grid, 118, 122
Regular lag interval classes, 168
Relative variogram, 137, 140
Residual sums of squares, 81, 88
Right-click menu, 23
Rodograms, 98
Rotate graph, 153, 155
Sample ID, 39, 167
Sample postings. *See Coordinate postings*
Sample variance, 74
Scaling data, 54
Screen layout, 15
Search neighborhood, 120
nearest neighbors, 120
octant search, 121
radius, 121
Secondary variate (Z2), 138
Seed value, 28
Semivariance
capacity, 72
overview, 70
semivariance analysis window, 71
semivariance offset, 70
values, 78
variograms window, 77
Serial number, 7, 8
Sill, 79, 80, 86, 87, 167
Simple cokriging, 140
Simple kriging, 137
Single-user license agreement, 12
Site licensing, 12
Smoothing factor, 28, 141, 142
Sort, 18
Space separated (free format) values, 52
Spaces as column title separators, 52
Specific shapes, 119
Spherical anisotropic model, 88
Spherical isotropic model, 82
Spreadsheet input files, 44, 51, 52
Standard deviations map, 154

Standardized ordinary cokriging, 140
Standardized variogram, 96
Structural variance, 79, 163, 167
Subsampled data, 27
Summary statistics, 54
Surface, 149
Surfer®
  grid output file format, 120
  map input files, 50, 145
  Map input files, 148
Surfer® XYZ input file format, 46
Symbols, 32, 66
Tab separated values, 52
Tabs as column title separators, 52
Text files, 44, 45
Tick marks, 31
Title, 38
Tool bar, 16
Transect image, 158
Transfer license, 22
Transformations
  assigning, 54
  backtransformations, 55, 162
  definition, 168
  lognormal, 54
  offsets, 55
  scaling 0-1, 54
  square root, 54
Trend, 99
Universal kriging, 136
Updates, 12, 22
Validation analysis, 121
  cross-validation analysis, 129
  jackknife analysis, 129
Variance
  cloud analysis, 77
  cloud pairs, 112
Variate assignments, 38
Variate names, 38, 47, 163, 164, 167, 168
Variogram
  A_0, 79, 80, 86, 87, 167
  active lag default, 27
  anisotropic models, 86
  anisotropy surface, 27
  C_0, 79, 80, 86, 87
  capacity, 72
  defaults, 27
  effective range, 79, 80, 86, 87, 167
  isotropic models, 79, 80
  map, 73, 92, 94
  model command, 74
  model parameters, 80, 168
  model type for kriging, 137, 140
  nugget variance, 79, 80, 86, 87
  number of lag intervals, 27
  offset, 27
  options, 74
  range, 79, 80, 86, 87, 167
  sill, 79, 80, 86, 87, 167
  structural variance, 79, 167
  subsample data, 27
  values, 78
  variograms window, 77
Variography, 70
Vertices, 126
Warnings, 24
Warranty, 13
Weighting power, 28, 141, 142
Window menu, 21
Worksheets
  capacity, 39
  formulas, 39
  jackknife data, 121
  limits, 39
  printing, 35
  sorting, 39
  window, 37
X axis range, 31
X-coordinate
  assignment, 39, 40
  defintion, 169
xls files, 18, 35
xml files, 18, 35
Y axis range, 31
Y-coordinate
  assignment, 39, 40

defintion, 169
Z (primary variate)
assignment, 39, 40
axis range, 32
definition, 169
estimate boundaries, 122
Z2 (covariate)
assignment, 39, 40
definition, 164, 169
icons, 16
regression, 67
values warning, 41
Zoom, 153